

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

AY

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup> :</b> <b>A61K 37/12, A61F 2/02</b> <b>C07K 13/00</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 92/15323</b> <b>(43) International Publication Date:</b> 17 September 1992 (17.09.92)
<b>(21) International Application Number:</b> PCT/US92/01968 <b>(22) International Filing Date:</b> 11 March 1992 (11.03.92)  <b>(30) Priority data:</b> 667,274 11 March 1991 (11.03.91) US  <b>(71) Applicant:</b> CREATIVE BIOMOLECULES, INC. [US/ US]; 35 South Street, Hopkinton, MA 01748 (US).  <b>(72) Inventors:</b> COHEN, Charles, M. ; 98 Winthrop Street, Medway, MA 02053 (US). KUBERASAMPATH, Than- gavel ; 6 Spring Street, Medway, MA 02053 (US). PANG, Roy, H., L. ; 16 Kimberly Drive, Medway, MA 02053 (US). OPPERMAN, Hermann ; 25 Summer Hill Road, Medway, MA 02053 (US). RUEGER, David, C. ; 19 Downey Street, Hopkinton, MA 01748 (US).		<b>(74) Agent:</b> PITCHER, Edmund, R.; Testa, Hurwitz & Thi- beault, Exchange Place, 53 State Street, Boston, MA 02109-2809 (US).  <b>(81) Designated States:</b> AT (European patent), AU, BE (Euro- pean patent), CA, CH (European patent), DE (Euro- pean patent), DK (European patent), ES (European pa- tent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (Euro- pean patent), MC (European patent), NL (European pa- tent), SE (European patent).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> PROTEIN-INDUCED MORPHOGENESIS  <b>(57) Abstract</b>  Disclosed are 1) amino acid sequence data, structural features, homologies and various other data characterizing morpho- genic proteins, 2) methods of producing these proteins from natural and recombinant sources and from synthetic constructs, 3) morphogenic devices comprising these morphogenic proteins and a suitably modified tissue-specific matrix, and 4) methods of inducing non-chondrogenic tissue growth in a mammal.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	MI	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NI	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

## PROTEIN-INDUCED MORPHOGENESIS

### Background of the Invention

This invention relates to morphogenic proteins which can induce tissue morphogenesis in mammals; to  
5 methods of identifying these proteins and obtaining them from natural sources or producing synthetic forms of these proteins by expressing recombinant DNA encoding the proteins; to the fabrication of tissue-specific acellular matrices; and to methods for  
10 promoting tissue stasis, repair and regeneration, and methods for increasing progenitor cell populations using these proteins.

Cell differentiation is the central  
15 characteristic of morphogenesis which initiates in the embryo, and continues to various degrees throughout the life of an organism in adult tissue repair and regeneration mechanisms. The degree of morphogenesis in adult tissue varies among different tissues and is  
20 related, among other things, to the degree of cell turnover in a given tissue. On this basis, tissues can be divided into three broad categories: (1) tissues with static cell populations such as nerve and skeletal muscle where there is no cell division and most of the

cells formed during early development persist throughout adult life; (2) tissues containing conditionally renewing populations such as liver where there is generally little cell division but, in response to an appropriate stimulus, cells can divide to produce daughters of the same differentially defined type; and (3) tissues with permanently renewing populations including blood, testes and stratified squamous epithelia which are characterized by rapid and continuous cell turnover in the adult. Here, the terminally differentiated cells have a relatively short life span and are replaced through proliferation of a distinct subpopulation of cells, known as stem or progenitor cells.

15

The cellular and molecular events which govern the stimulus for differentiation of these cells is an area of intensive research. In the medical field, it is anticipated that the discovery of factor(s) which control cell differentiation and tissue morphogenesis will significantly advance medicine's ability to repair and regenerate diseased or damaged mammalian tissues and organs. Particularly useful areas include reconstructive surgery and in the treatment of tissue degenerative diseases including arthritis, emphysema, osteoporosis, cardiomyopathy, cirrhosis, and degenerative nerve diseases.

A number of different factors have been isolated in recent years which appear to play a role in cell differentiation. Some of these factors are gene transcription activators such as the NOTCH gene, identified in *Drosophila* and the related XOTCH gene identified in *Xenopus*, as well as a number of transcription activators identified in *Caenorhabditis elegans*.

The hemopoietic system, because of its continually renewing cell population, is an area of concentrated study. Factors identified in this system which may be involved in cell renewal include

- 5 interleukin 3 (IL-3), erythropoietin, the CSFs (GM-CSF, G-CSF, M-CSF et al.) and various stem cell growth factors.

- Other proteins thought to play a role in cell differentiation include proteins that are members of the family of insulin-like growth factors (IGF), members of the family of heparin-binding growth factors, (e.g., FGF - acidic and basic fibroblast growth factors, and ECDGF - embryonal carcinoma-derived growth factor) as well as several transforming oncogenes (hst and int-2, see for example, Heath et al., (1988), J. Cell Sci. Suppl. 10:256-256.) DIF (Differentiation Inducing Factor), identified in Dictyostelium discoideum, is another bioregulatory protein, directing prestock cell differentiation in that organism.

- The structurally related proteins of the TGF- $\beta$  superfamily of proteins also have been identified as involved in a variety of developmental events. For example, TGF- $\beta$  and the polypeptides of the inhibin/activin group appear to play a role in the regulation of cell growth and differentiation. MIS (Mullerian Inhibiting Substance) causes regression of the Mullerian duct in development of the mammalian male embryo, and DPP, the gene product of the *Drosophila* decapentaplegic complex is required for appropriate dorsal-ventral specification. Similarly, Vg-1 is involved in mesoderm induction in *Xenopus*, and Vgr-1 has been identified in a variety of developing murine tissues.

Another source that has revealed a wealth of information is in the area of bone morphogenesis. The development and study of a bone model system has  
5 identified the developmental cascade of bone differentiation as consisting of chemotaxis of mesenchymal cells, proliferation of these progenitor cells, differentiation of these cells into chondroblasts, cartilage calcification, vascular  
10 invasion, bone formation, remodeling, and finally, marrow differentiation (Reddi (1981) Collagen Rel. Res. 1:209-206). Proteins capable of inducing endochondral bone formation in a mammal when implanted in association with a matrix now have been identified in a  
15 number of different mammalian species, as have the genes encoding these proteins, (see, for example, U.S. Patent No. 4,968,590 and U.S. Patent No. 5,011,691, Ozkaynak, et al., (1990) EMBO J 9:2085-2093, and Ozkaynak et al., (1991) Biochem. Biophys. Res. Commun.  
20 179:116-123 and USSN 07/841,646, filed February 21, 1992.) These proteins, which share significant amino acid sequence homology with one another as well as structural similarities with various members of the TGF- $\beta$  super family of proteins, have been shown to  
25 induce endochondral bone formation and/or cartilage formation when implanted in a mammal in association with a suitably modified matrix. Proteins capable of inducing a similar developmental cascade of tissue morphogenesis of other tissues have not been  
30 identified.

It is an object of this invention to provide morphogenic proteins ("morphogens"), and methods for identifying these proteins, which are capable of  
35 inducing the developmental cascade of tissue

morphogenesis for a variety of tissues in mammals different from bone or cartilage. This morphogenic activity includes the ability to induce proliferation and differentiation of progenitor cells, and the

5 ability to support and maintain the differentiated phenotype through the progression of events that results in the formation of adult tissue. Another object is to provide genes encoding these proteins as well as methods for the expression and isolation of

10 these proteins, from either natural sources or biosynthetic sources, using recombinant DNA techniques. Still another object is to provide tissue-specific acellular matrices that may be used in combination with these proteins, and methods for their production.

15 Other objects include providing methods for increasing a progenitor cell population in a mammal, methods for stimulating progenitor cells to differentiate in vivo or in vitro and maintain their differentiated phenotype, methods for inducing tissue-specific growth

20 in vivo and methods for the replacement of diseased or damaged tissue in vivo. These and other objects and features of the invention will be apparent from the description, drawings, and claims which follow.

Summary of the Invention

This invention provides morphogenic proteins ("morphogens") capable of inducing the developmental cascade of tissue morphogenesis in a mammal. In particular, these proteins are capable of inducing the proliferation of uncommitted progenitor cells, and inducing the differentiation of these stimulated progenitor cells in a tissue-specific manner under appropriate environmental conditions. In addition, the morphogens are capable of supporting the growth and maintenance of these differentiated cells. These morphogenic activities allow the proteins of this invention to initiate and maintain the developmental cascade of tissue morphogenesis in an appropriate, morphogenically permissive environment, stimulating stem cells to proliferate and differentiate in a tissue-specific manner, and inducing the progression of events that culminate in new tissue formation. These morphogenic activities also allow the proteins to stimulate the "redifferentiation" of cells previously induced to stray from their differentiation path. Under appropriate environmental conditions it is anticipated that these morphogens also may stimulate the "dedifferentiation" of committed cells (see infra.)

In one aspect of the invention, the proteins and compositions of this invention are useful in the replacement of diseased or damaged tissue in a mammal, particularly when the damaged tissue interferes with normal tissue or organ function. Accordingly, it is anticipated that the proteins of this invention will be useful in the repair of damaged tissue such as, for example, damaged lung tissue resulting from emphysema, cirrhotic kidney or liver tissue, damaged heart or



blood vessel tissue, as may result from cardiomyopathies and/or atherothrombotic or cardioembolic strokes, damaged stomach tissue resulting from ulceric perforations or their repair, damaged  
5 neural tissue as may result from physical injury, degenerative diseases such as Alzheimer's disease or multiple sclerosis or strokes, damaged dentin tissue as may result from disease or mechanical injury. When the proteins of this invention are provided to, or their  
10 expression stimulated at, a tissue-specific locus, the developmental cascade of tissue morphogenesis is induced (see infra). Cells stimulated ex vivo by contact with the proteins or agents capable of stimulating morphogen expression in these cells also  
15 may be provided to the tissue locus. In these cases the existing tissue provides the necessary matrix requirements, providing a suitable substratum for the proliferating and differentiating cells in a morphogenically permissive environment, as well as  
20 providing the necessary signals for directing the tissue-specificity of the developing tissue. Alternatively, the proteins or stimulated cells may be combined with a formulated matrix and implanted as a device at a locus in vivo. The formulated matrix  
25 should be a biocompatible, preferably biodegradable, appropriately modified tissue-specific acellular matrix having the characteristics described below.

In many instances, the loss of tissue function  
30 results from scar tissue, formed in response to an initial or repeated injury to the tissue. The degree of scar tissue formation generally depends on the regenerative properties of the injured tissue, and on the degree and type of injury. Thus, in another

aspect, the invention includes morphogens that may be used to prevent or substantially inhibit the formation of scar tissue by providing the morphogens, or morphogen-stimulated cells, to a newly injured tissue  
5 loci (see infra).

The morphogens of this invention also may be used to increase or regenerate a progenitor or stem cell population in a mammal. For example, progenitor  
10 cells may be isolated from an individual's bone marrow, stimulated ex vivo for a time and at a morphogen concentration sufficient to induce the cells to proliferate, and returned to the bone marrow. Other sources of progenitor cells that may be suitable  
15 include biocompatible cells obtained from a cultured cell line, stimulated in culture, and subsequently provided to the body. Alternatively, the morphogen may be provided systemically, or implanted, injected or otherwise provided to a progenitor cell population in  
20 an individual to induce its mitogenic activity in vivo. For example, an agent capable of stimulating morphogen expression in the progenitor cell population of interest may be provided to the cells in vivo, for example systemically, to induce mitogenic activity.  
25 Similarly, a particular population of hemopoietic stem cells may be increased by the morphogens of this invention, for example by perfusing an individual's blood to extract the cells of interest, stimulating these cells ex vivo, and returning the stimulated cells  
30 to the blood. It is anticipated that the ability to augment an individual's progenitor cell population will significantly enhance existing methods for treating disorders resulting from a loss or reduction of a renewable cell population. Two particularly  
35 significant applications include the treatment of blood

disorders and impaired or lost immune function. Other cell populations whose proliferation may be exploited include the stem cells of the epidermis, which may be used in skin tissue regeneration, and the stem cells of the gastrointestinal lining, for example, in the healing of ulcers.

In still another aspect of the invention, the morphogens also may be used to support the growth and maintenance of differentiated cells, inducing existing differentiated cells to continue expressing their phenotype. It is anticipated that this activity will be particularly useful in the treatment of tissue disorders where loss of function is caused by cells becoming senescent or quiescent, such as may occur in osteoporosis. Application of the protein directly to the cells to be treated, or providing it by systemic injection, can be used to stimulate these cells to continue expressing their phenotype, thereby significantly reversing the effects of the dysfunction (see *infra*). Alternatively, administration of an agent capable of stimulating morphogen expression in vivo also may be used. In addition, the morphogens of this invention also may be used in gene therapy protocols to stimulate the growth of quiescent cells, thereby potentially enhancing the ability of these cells to incorporate exogenous DNA.

In yet another aspect of the invention, the morphogens of this invention also may be used to induce "redifferentiation" of cells that have strayed from their differentiation pathway, such as can occur during tumorigenesis. It is anticipated that this activity of the proteins will be particularly useful in treatments to reduce or substantially inhibit the growth of

neoplasms. The method also is anticipated to induce the de-and re-differentiation of these cells. As described supra, the proteins may be provided to the cells directly or systemically, or an agent capable of stimulating morphogen expression in vivo may be provided.

Finally, modulations of endogenous morphogen levels may be monitored as part of a method for detecting tissue dysfunction. Specifically, modulations in endogenous morphogen levels are anticipated to reflect changes in tissue or organ stasis. Tissue stasis may be monitored by detecting changes in the levels of the morphogen itself. For example, tissue samples may be obtained at intervals and the concentration of the morphogen present in the tissue detected by standard protein detection means known to those skilled in the art. As an example, a binding protein capable of interacting specifically with the morphogen of interest, such as an anti-morphogen antibody, may be used to detect the morphogen in a standard immunoassay. The morphogen levels detected then may be compared, the changes in the detected levels being indicative of the status of the tissue. Modulations in endogenous morphogen levels also may be monitored by detecting fluctuations in the body's natural antibody titer to morphogens (see infra.)

The morphogenic proteins and compositions of this invention can be isolated from a variety of naturally-occurring sources, or they may be constructed biosynthetically using conventional recombinant DNA technology. Similarly, the matrices may be derived from organ-specific tissue, or they may be formulated synthetically, as described below.

A key to these developments was the discovery and characterization of naturally-occurring osteogenic proteins followed by observation of their remarkable properties. These proteins, originally isolated from bone, are capable of inducing the full developmental cascade of bone formation, including vascularization, mineralization, and bone marrow differentiation, when implanted in a mammalian body in association with a suitably modified matrix. Native proteins capable of inducing this developmental cascade, as well as DNA sequences encoding these proteins now have been isolated and characterized for a number of different species (e.g., human and mouse OP-1, OP-2, and CBMP-2. See, for example, U.S. Patent Nos. 4,968,590 and 5,011,691; U.S. Application Serial No. 841,646, filed February 21, 1992; Sampath et al. (1990) J. Bio. Chem 265:13198-13205; Ozkaynak, et al. (1990) EMBO J 9:2085-2 093 and Ozkaynak, et al. (1991) Biochem. Biophys. Res. Commn. 179:116-123.) The mature forms of these proteins share substantial amino acid sequence homology, especially in the C-terminal regions of the mature proteins. In particular, the proteins share a conserved six or seven cysteine skeleton in this region (e.g., the linear arrangement of these C-terminal cysteine residues is essentially conserved in the different proteins, in addition to other, apparently required amino acids (see Table II, *infra*)).

Polypeptide chains not normally associated with bone or bone formation, but sharing substantial amino acid sequence homology with the C-terminus of the osteogenic proteins, including the conserved six or seven cysteine skeleton, also have been identified as competent for inducing bone in mammals. Among these are amino acid sequences identified in *Drosophila* and

Xenopus, (e.g., DPP and Vgl; see, for example, U.S. Patent No. 5,011,691 and Table II, *infra*). In addition, non-native biosynthetic constructs designed based on extrapolation from these sequence homologies, including the conserved six or seven cysteine skeleton, have been shown to induce endochondral bone formation in mammals when implanted in association with an appropriate matrix (see U.S. Pat. No. 5,011,691 and Table III, *infra*).

10

It has now been discovered that this "family" of proteins sharing substantial amino acid sequence homology and the conserved six or seven cysteine skeleton are true morphogens, capable of inducing, in addition to bone and cartilage, tissue-specific morphogenesis for a variety of other organs and tissues. The proteins apparently bind to surface receptors or otherwise contact and interact with progenitor cells, predisposing or stimulating the cells to proliferate and differentiate in a morphogenically permissive environment. The morphogens are capable of inducing the developmental cascade of cellular and molecular events that culminate in the formation of new organ-specific tissue, including any vascularization, connective tissue formation, and nerve enervation as required by the naturally occurring tissue.

It also has been discovered that the way in which the cells differentiate, whether, for example, they differentiate into bone-producing osteoblasts, hemopoietic cells, or liver cells, depends on the nature of their local environment (see *infra*). Thus, in addition to requiring a suitable substratum on which to anchor, the proliferating and differentiating cells also require appropriate signals to direct their

35

tissue-specificity. These signals may take the form of cell surface markers.

<

5           When the morphogens (or progenitor cells stimulated by these morphogens) are provided at a tissue-specific locus (e.g., by systemic injection or by implantation or injection at a tissue-specific locus, or by administration of an agent capable of  
10 stimulating morphogen expression in vivo), the existing tissue at that locus, whether diseased or damaged, has the capacity of acting as a suitable matrix. Alternatively, a formulated matrix may be externally provided together with the stimulated progenitor cells  
15 or morphogen, as may be necessary when the extent of injury sustained by the damaged tissue is large. The matrix should be a biocompatible, suitably modified acellular matrix having dimensions such that it allows the influx, differentiation, and proliferation of  
20 migratory progenitor cells, and is capable of providing a morphogenically permissive environment (see *infra*). The matrix preferably is tissue-specific, and biodegradable.

25           Formulated matrices may be generated from dehydrated organ-specific tissue, prepared for example, by treating the tissue with solvents to substantially remove the non-structural components from the tissue. Alternatively, the matrix may be formulated  
30 synthetically using a biocompatible, preferably in vivo

biodegradable, structural polymer such as collagen in association with suitable tissue-specific cell attachment factors. Currently preferred structural polymers comprise tissue-specific collagens. Currently preferred cell attachment factors include glycosaminoglycans and proteoglycans. The matrix further may be treated with an agent or agents to increase the number of pores and micropits on its surfaces, so as to enhance the influx, proliferation and differentiation of migratory progenitor cells from the body of the mammal.

Among the proteins useful in this invention are proteins originally identified as osteogenic proteins, such as the OP-1, OP-2 and CBMP2 proteins, as well as amino acid sequence-related proteins such as DPP (from *Drosophila*), Vgl (from *Xenopus*), Vgr-1 (from mouse, see Table II and Seq. ID Nos. 5-14), and the recently identified GDF-1 protein (Seq. ID No. 14). The members of this family, which include members of the TGF- $\beta$  super-family of proteins, share substantial amino acid sequence homology in their C-terminal regions. Table I, below, describes the various morphogens identified to date, including their nomenclature as used herein, and Seq. ID references.

TABLE I

30	"OP-1"	Refers generically to the group of morphogenically active proteins expressed from part or all of a DNA sequence encoding OP-1 protein, including allelic and species variants thereof, e.g., human OP-1 ("hOP-1", Seq. ID No. 5, mature
35		



protein amino acid sequence), or mouse  
OP-1 ("mOP-1", Seq. ID No. 6, mature  
protein amino acid sequence.) The  
conserved seven cysteine skeleton is  
5 defined by residues 38 to 139 of Seq. ID  
Nos. 5 and 6. The cDNA sequences and the  
amino acids encoding the full length  
proteins are provided in Seq. Id Nos. 16  
and 17 (hOP1) and Seq. ID Nos. 18 and 19  
10 (mOP1.) The mature proteins are defined  
by residues 293-431 (hOP1) and 292-430  
(mOP1). The "pro"regions of the proteins,  
cleaved to yield the mature,  
morphogenically active proteins are  
15 defined essentially by residues 30-292  
(hOP1) and residues 30-291 (mOP1).

"OP-2"

20 refers generically to the group of active  
proteins expressed from part or all of a  
DNA sequence encoding OP-2 protein,  
including allelic and species variants  
thereof, e.g., human OP-2 ("hOP-2", Seq.  
ID No. 7, mature protein amino acid  
sequence) or mouse OP-2 ("mOP-2", Seq. ID  
25 No. 8, mature protein amino acid  
sequence). The conserved seven cysteine  
skeleton is defined by residues 38 to 139  
of Seq. ID Nos. 7 and 8. The cDNA  
sequences and the amino acids encoding the  
30 full length proteins are provided in Seq.  
Id Nos. 20 and 21 (hOP2) and Seq. ID Nos.  
22 and 23 (mOP2.) The mature proteins are  
defined essentially by residues 264-402  
(hOP2) and 261-399 (mOP2). The "pro"  
35 regions of the proteins, cleaved to yield

the mature, morphogenically active proteins are defined essentially by residues 18-263 (hOP2) and residues 18-260 (mOP1).

- 5  
"CBMP2" refers generically to the morphogenically active proteins expressed from a DNA sequence encoding the CBMP2 proteins, including allelic and species variants thereof, e.g., human CBMP2A ("CBMP2A(fx)", Seq ID No. 9) or human CBMP2B DNA ("CBMP2B(fx)", Seq. ID No. 10).
- 10
- 15 "DPP(fx)" refers to protein sequences encoded by the Drosophila DPP gene and defining the conserved seven cysteine skeleton (seq. ID No. 11).
- 20 "Vgl(fx)" refers to protein sequences encoded by the Xenopus Vgl gene and defining the conserved seven cysteine skeleton (Seq. ID No. 12).
- 25 "Vgr-1(fx)" refers to protein sequences encoded by the murine Vgr-1 gene and defining the conserved seven cysteine skeleton (Seq. ID No. 13).
- 30 "GDF-1(fx)" refers to protein sequences encoded by the human GDF-1 gene and defining the conserved seven cysteine skeleton (seq. ID No. 14).

The OP-2 proteins have an additional cysteine residue in this region (e.g., see residue 41 of Seq. ID Nos. 7 and 8), in addition to the conserved cysteine skeleton in common with the other proteins in this family. The GDF-1 protein has a four amino acid insert within the conserved skeleton (residues 44-47 of Seq. ID No. 14) but this insert likely does not interfere with the relationship of the cysteines in the folded structure. In addition, the CBMP2 proteins are missing one amino acid residue within the cysteine skeleton.

The morphogens are inactive when reduced, but are active as oxidized homodimers and when oxidized in combination with other morphogens of this invention. Thus, as defined herein, a morphogen of this invention is a dimeric protein comprising a pair of polypeptide chains, wherein each polypeptide chain comprises at least the C-terminal six cysteine skeleton defined by residues 43-139 of Seq. ID No. 5, including functionally equivalent arrangements of these cysteines (e.g., amino acid insertions or deletions which alter the linear arrangement of the cysteines in the sequence but not their relationship in the folded structure), such that, when the polypeptide chains are folded, the dimeric protein species comprising the pair of polypeptide chains has the appropriate three-dimensional structure, including the appropriate intra- or inter-chain disulfide bonds such that the protein is capable of acting as a morphogen as defined herein. Specifically, the protein is capable of any of the following biological functions in a morphogenically permissive environment: stimulating proliferation of progenitor cells; stimulating the differentiation of progenitor cells; stimulating the proliferation of differentiated cells; and supporting the growth and maintenance of differentiated cells, including the

"redifferentiation" of these cells. In addition, it is also anticipated that the morphogens of this invention will be capable of inducing dedifferentiation of committed cells under appropriate environmental  
5 conditions.

In one preferred aspect, the morphogens of this invention comprise one of two species of generic amino acid sequences: Generic Sequence 1 (Seq. ID  
10 No. 1) or Generic Sequence 2 (Seq. ID No. 2); where each Xaa indicates one of the 20 naturally-occurring L-isomer,  $\alpha$ -amino acids or a derivative thereof. Generic Sequence 1 comprises the conserved six cysteine skeleton and Generic Sequence 2 comprises the conserved  
15 six cysteine skeleton plus the additional cysteine identified in OP-2 (see residue 36, Seq. ID No. 2). In another preferred aspect, these sequences further comprise the following additional sequence at their N-terminus:

20

Cys Xaa Xaa Xaa Xaa (Seq. ID No. 15)

1

5

Preferred amino acid sequences within the  
25 foregoing generic sequences include: Generic Sequence 3 (Seq. ID No. 3) and Generic Sequence 4 (Seq. ID No. 4), listed below, which accommodate the homologies shared among the various preferred members of this morphogen family identified to date (see Table II), as  
30 well as the amino acid sequence variation among them. Generic Sequences 3 and 4 are composite amino acid sequences of the proteins presented in Table II and identified in Seq. ID Nos. 5-14. The generic sequences include both the amino acid identity shared by the  
35 sequences in Table II, as well as alternative residues

for the variable positions within the sequence. Note that these generic sequences allow for an additional cysteine at position 41 or 46 in Generic Sequences 3 or 4, respectively, providing an appropriate cysteine skeleton where inter- or intramolecular disulfide bonds can form, and contain certain critical amino acids which influence the tertiary structure of the proteins.

### Generic Sequence 3

10		Leu	Tyr	Val	Xaa	Phe	
		1				5	
		Xaa	Xaa	Xaa	Gly	Trp	Xaa Xaa Trp Xaa
					10		
		Xaa	Ala	Pro	Xaa	Gly	Xaa Xaa Ala
15		15				20	
		Xaa	Tyr	Cys	Xaa	Gly	Xaa Cys Xaa
				25			30
		Xaa	Pro	Xaa	Xaa	Xaa	Xaa Xaa
					35		
20		Xaa	Xaa	Xaa	Asn	His	Ala Xaa Xaa
				40			45
		Xaa	Xaa	Leu	Xaa	Xaa	Xaa Xaa Xaa
					50		
		Xaa	Xaa	Xaa	Xaa	Xaa	Xaa Xaa Cys
25				55			60
		Cys	Xaa	Pro	Xaa	Xaa	Xaa Xaa Xaa
					65		

Xaa Xaa Xaa Leu Xaa Xaa Xaa

70

75

Xaa Xaa Xaa Xaa Val Xaa Leu Xaa

80

5 Xaa Xaa Xaa Xaa Met Xaa Val Xaa

85

90

Xaa Cys Gly Cys Xaa

95

wherein each Xaa is independently selected from a group  
10 of one or more specified amino acids defined as  
follows: "Res." means "residue" and Xaa at res.4 =  
(Ser, Asp or Glu); Xaa at res.6 = (Arg, Gln, Ser or  
Lys); Xaa at res.7 = (Asp or Glu); Xaa at res.8 = (Leu  
or Val); Xaa at res.11 = (Gln, Leu, Asp, His or Asn);  
15 Xaa at res.12 = (Asp, Arg or Asn); Xaa at res.14 = (Ile  
or Val); Xaa at res.15 = (Ile or Val); Xaa at res.18 =  
(Glu, Gln, Leu, Lys, Pro or Arg); Xaa at res.20 = (Tyr  
or Phe); Xaa at res.21 = (Ala, Ser, Asp, Met, His, Leu  
or Gln); Xaa at res.23 = (Tyr, Asn or Phe); Xaa at  
20 res.26 = (Glu, His, Tyr, Asp or Gln); Xaa at res.28 =  
(Glu, Lys, Asp or Gln); Xaa at res.30 = (Ala, Ser, Pro  
or Gln); Xaa at res.31 = (Phe, Leu or Tyr); Xaa at  
res.33 = (Leu or Val); Xaa at res.34 = (Asn, Asp, Ala  
or Thr); Xaa at res.35 = (Ser, Asp, Glu, Leu or Ala);  
25 Xaa at res.36 = (Tyr, Cys, His, Ser or Ile); Xaa at  
res.37 = (Met, Phe, Gly or Leu); Xaa at res.38 = (Asn  
or Ser); Xaa at res.39 = (Ala, Ser or Gly); Xaa at  
res.40 = (Thr, Leu or Ser); Xaa at res.44 = (Ile or  
Val); Xaa at res.45 = (Val or Leu); Xaa at res.46 =  
30 (Gln or Arg); Xaa at res.47 = (Thr, Ala or Ser); Xaa at  
res.49 = (Val or Met); Xaa at res.50 = (His or Asn);  
Xaa at res.51 = (Phe, Leu, Asn, Ser, Ala or Val); Xaa

at res.52 = (Ile, Met, Asn, Ala or Val); Xaa at res.53  
 = (Asn, Lys, Ala or Glu); Xaa at res.54 = (Pro or Ser);  
 Xaa at res.55 = (Glu, Asp, Asn, or Gly); Xaa at res.56  
 = (Thr, Ala, Val, Lys, Asp, Tyr, Ser or Ala); Xaa at  
 5 res.57 = (Val, Ala or Ile); Xaa at res.58 = (Pro or  
 Asp); Xaa at res.59 = (Lys or Leu); Xaa at res.60 =  
 (Pro or Ala); Xaa at res.63 = (Ala or Val); Xaa at  
 res.65 = (Thr or Ala); Xaa at res.66 = (Gln, Lys, Arg  
 or Glu); Xaa at res.67 = (Leu, Met or Val); Xaa at  
 10 res.68 = (Asn, Ser or Asp); Xaa at res.69 = (Ala, Pro  
 or Ser); Xaa at res.70 = (Ile, Thr or Val); Xaa at  
 res.71 = (Ser or Ala); Xaa at res.72 = (Val or Met);  
 Xaa at res.74 = (Tyr or Phe); Xaa at res.75 = (Phe, Tyr  
 or Leu); Xaa at res.76 = (Asp or Asn); Xaa at res.77 =  
 15 (Asp, Glu, Asn or Ser); Xaa at res.78 = (Ser, Gln, Asn  
 or Tyr); Xaa at res.79 = (Ser, Asn, Asp or Glu); Xaa at  
 res.80 = (Asn, Thr or Lys); Xaa at res.82 = (Ile or  
 Val); Xaa at res.84 = (Lys or Arg); Xaa at res.85 =  
 (Lys, Asn, Gln or His); Xaa at res.86 = (Tyr or His);  
 20 Xaa at res.87 = (Arg, Gln or Glu); Xaa at res.88 =  
 (Asn, Glu or Asp); Xaa at res.90 = (Val, Thr or Ala);  
 Xaa at res.92 = (Arg, Lys, Val, Asp or Glu); Xaa at  
 res.93 = (Ala, Gly or Glu); and Xaa at res.97 = (His or  
 Arg); and Generic Seq. 4:

25

Generic Sequence 4

	Cys	Xaa	Xaa	Xaa	Xaa	Leu	Tyr	Val	Xaa	Phe
	1				5					10
30	Xaa	Xaa	Xaa	Gly	Trp	Xaa	Xaa	Trp	Xaa	
				15						
	Xaa	Ala	Pro	Xaa	Gly	Xaa	Xaa	Ala		
	20				25					
	Xaa	Tyr	Cys	Xaa	Gly	Xaa	Cys	Xaa		
35			30					35		

```

      Xaa Pro Xaa Xaa Xaa Xaa Xaa
                40
      Xaa Xaa Xaa Asn His Ala Xaa Xaa
                45                      50
5      Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa
                55
      Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
                60                      65
      Cys Xaa Pro Xaa Xaa Xaa Xaa Xaa
10      70
      Xaa Xaa Xaa Leu Xaa Xaa Xaa
                75                      80
      Xaa Xaa Xaa Xaa Val Xaa Leu Xaa
                85
15      Xaa Xaa Xaa Xaa Met Xaa Val Xaa
                90                      95
      Xaa Cys Gly Cys Xaa
                100

```

wherein each Xaa is independently selected from a group  
 20 of one or more specified amino acids as defined by the  
 following: "Res." means "residue" and Xaa at res.2 =  
 (Lys or Arg); Xaa at res.3 = (Lys or Arg); Xaa at res.4  
 = (His or Arg); Xaa at res.5 = (Glu, Ser, His, Gly, Arg  
 or Pro); Xaa at res.9 = (Ser, Asp or Glu); Xaa at  
 25 res.11 = (Arg, Gln, Ser or Lys); Xaa at res.12 = (Asp  
 or Glu); Xaa at res.13 = (Leu or Val); Xaa at res.16 =  
 (Gln, Leu, Asp, His or Asn); Xaa at res.17 = (Asp, Arg,  
 or Asn); Xaa at res.19 = (Ile or Val); Xaa at res.20 =  
 (Ile or Val); Xaa at res.23 = (Glu, Gln, Leu, Lys, Pro  
 30 or Arg); Xaa at res.25 = (Tyr or Phe); Xaa at res.26 =  
 (Ala, Ser, Asp, Met, His, Leu, or Gln); Xaa at res.28 =  
 (Tyr, Asn or Phe); Xaa at res.31 = (Glu, His, Tyr, Asp  
 or Gln); Xaa at res.33 = Glu, Lys, Asp or Gln); Xaa at  
 res.35 = (Ala, Ser or Pro); Xaa at res.36 = (Phe, Leu  
 35 or Tyr); Xaa at res.38 = (Leu or Val); Xaa at res.39 =



(Asn, Asp, Ala or Thr); Xaa at res.40 = (Ser, Asp, Glu, Leu or Ala); Xaa at res.41 = (Tyr, Cys, His, Ser or Ile); Xaa at res.42 = (Met, Phe, Gly or Leu); Xaa at res.44 = (Ala, Ser or Gly); Xaa at res.45 = (Thr, Leu or Ser); Xaa at res.49 = (Ile or Val); Xaa at res.50 = (Val or Leu); Xaa at res.51 = (Gln or Arg); Xaa at res.52 = (Thr, Ala or Ser); Xaa at res.54 = (Val or Met); Xaa at res.55 = (His or Asn); Xaa at res.56 = (Phe, Leu, Asn, Ser, Ala or Val); Xaa at res.57 = (Ile, Met, Asn, Ala or Val); Xaa at res.58 = (Asn, Lys, Ala or Glu); Xaa at res.59 = (Pro or Ser); Xaa at res.60 = (Glu, Asp, or Gly); Xaa at res.61 = (Thr, Ala, Val, Lys, Asp, Tyr, Ser or Ala); Xaa at res.62 = (Val, Ala or Ile); Xaa at res.63 = (Pro or Asp); Xaa at res.64 = (Lys or Leu); Xaa at res.65 = (Pro or Ala); Xaa at res.68 = (Ala or Val); Xaa at res.70 = (Thr or Ala); Xaa at res.71 = (Gln, Lys, Arg or Glu); Xaa at res.72 = (Leu, Met or Val); Xaa at res.73 = (Asn, Ser or Asp); Xaa at res.74 = (Ala, Pro or Ser); Xaa at res.75 = (Ile, Thr or Val); Xaa at res.76 = (Ser or Ala); Xaa at res.77 = (Val or Met); Xaa at res.79 = (Tyr or Phe); Xaa at res.80 = (Phe, Tyr or Leu); Xaa at res.81 = (Asp or Asn); Xaa at res.82 = (Asp, Glu, Asn or Ser); Xaa at res.83 = (Ser, Gln, Asn or Tyr); Xaa at res.84 = (Ser, Asn, Asp or Glu); Xaa at res.85 = (Asn, Thr or Lys); Xaa at res.87 = (Ile or Val); Xaa at res.89 = (Lys or Arg); Xaa at res.90 = (Lys, Asn, Gln or His); Xaa at res.91 = (Tyr or His); Xaa at res.92 = (Arg, Gln or Glu); Xaa at res.93 = (Asn, Glu or Asp); Xaa at res.95 = (Val, Thr or Ala); Xaa at res.97 = (Arg, Lys, Val, Asp or Glu); Xaa at res.98 = (Ala, Gly or Glu); and Xaa at res.102 = (His or Arg).

Particularly useful sequences for use as morphogens in this invention include the C-terminal

domains, e.g., the C-terminal 96-102 amino acid residues of Vgl, Vgr-1, DPP, OP-1, OP-2, CBMP-2A, CBMP-2B and GDF-1 (see Table II, *infra*, and Seq. ID Nos. 5-14) which include at least the conserved six or  
5 seven cysteine skeleton. In addition, biosynthetic constructs designed from the generic sequences, such as COP-1, 3-5, 7, 16 (see Table III, *infra*) also are useful. Other sequences include the C-terminal CBMP3 and the inhibins/activin proteins (see, for example,  
10 U.S. Pat. Nos. 4,968,590 and 5,011,691). Accordingly, other useful sequences are those sharing at least 70% amino acid sequence homology, and preferably 80% homology with any of the sequences above. These are anticipated to include allelic and species variants and  
15 mutants, and biosynthetic muteins, as well as novel members of this morphogenic family of proteins. Particularly envisioned in the family of related proteins are those proteins exhibiting morphogenic activity and wherein the amino acid changes from the  
20 preferred sequences include conservative changes, e.g., those as defined by Dayoff et al., Atlas of Protein Sequence and Structure; vol. 5, Suppl. 3, pp. 345-362, (M.O. Dayoff, ed., Nat'l BioMed. Research Fdn., Washington, D.C. 1979).

25

The currently most preferred protein sequences useful as morphogens in this invention include those having greater than 60% identity, preferably greater than 65% identity, with the amino acid sequence  
30 defining the conserved six cysteine skeleton of hOP1 (e.g., residues 43-139 of Seq. ID No. 5). These most preferred sequences include both allelic and species variants of the OP1 and OP2 proteins.

The invention thus provides proteins comprising any of the polypeptide chains described above, whether isolated from naturally-occurring sources, or produced by recombinant DNA techniques, and  
5 includes allelic and species variants of these proteins, naturally-occurring or biosynthetic mutants thereof, as well as various truncated and fusion constructs. Deletion or addition mutants also are envisioned to be active (see infra), including those  
10 which may alter the conserved C-terminal cysteine skeleton, provided that the alteration does not functionally disrupt the relationship of these cysteines in the folded structure. Accordingly, such active forms are considered the equivalent of the  
15 specifically described constructs disclosed herein. The proteins may include forms having varying glycosylation patterns, varying N-termini, a family of related proteins having regions of amino acid sequence homology, and active truncated or mutated forms of  
20 native or biosynthetic proteins, produced by expression of recombinant DNA in host cells.

The morphogenic proteins can be expressed from intact or truncated cDNA or from synthetic DNAs in  
25 procaryotic or eucaryotic host cells, and purified, cleaved, refolded, and dimerized to form morphogenically active compositions. Currently preferred host cells include E. coli or mammalian cells, such as CHO, COS or BSC cells.

30

Thus, in view of this disclosure, skilled genetic engineers can isolate genes from cDNA or genomic libraries of various different species which encode appropriate amino acid sequences, or construct  
35 DNAs from oligonucleotides, and then can express them

in various types of host cells, including both procaryotes and eucaryotes, to produce large quantities of active proteins capable of inducing tissue-specific cell differentiation and tissue morphogenesis in a  
5 variety of mammals including humans.

The invention thus further comprises these methods of inducing tissue-specific morphogenesis using the morphogenic proteins of this invention and  
10 pharmaceutical and therapeutic agents comprising the morphogens of this invention. The invention further comprises the use of these morphogens in the manufacture of pharmaceuticals for various medical procedures, including procedures for inducing tissue  
15 growth, procedures for inducing progenitor cell proliferation, procedures to inhibit neoplasm growth and procedures to promote phenotypic cell expression of differentiated cells.

Brief Description of the Drawings

The foregoing and other objects and features of this invention, as well as the invention itself, may  
5 be more fully understood from the following description, when read together with the accompanying drawings, in which:

FIGURE 1 is a photomicrograph of a Northern  
10 Blot identifying Vgr-1 specific transcripts in various adult murine tissues;

FIGURE 2 is a photomicrograph of a Northern  
Blot identifying mOP-1-specific mRNA expression in  
15 various murine tissues prepared from 2 week old mice (panel A) and 5 week old mice (Panel B);

FIGURE 3 is a photomicrograph of Northern  
Blots identifying mRNA expression of EF-Tu  
20 (A, control), mOP-1 (B, D), and Vgr-1 (C) in (1) 17-day embryos and (2) 3-day post natal mice;

FIGURE 4A and 4B are photomicrographs showing  
the presence of OP-1 (by immunofluorescence staining)  
25 in the cerebral cortex (A) and spinal cord (B);

FIGURE 5A and 5B are photomicrographs  
illustrating the ability of morphogen (OP-1) to induce  
undifferentiated NG108 cells (5A) to undergo  
30 differentiation of neural morphology (5B).

FIGURE 6A-6D are photomicrographs showing the  
effect of morphogen (OP-1) on human embryo carcinoma  
cell redifferentiation;

FIGURE 7 is a photomicrograph showing the effects of phosphate buffered saline (PBS, animal 1) or morphogen (OP-1, animal 2) on partially hepatectomized rats;

5

FIGURE 8A - 8C are photomicrographs showing the effect of no treatment (8A), carrier matrix treatment (8B) and morphogen treatment (OP-1,8C) on dentin regeneration.

10

Detailed Description

Purification protocols first were developed which enabled isolation of the osteogenic (bone inductive) protein present in crude protein extracts from mammalian bone. (See PCT US 89/01453, and U.S. 4,968,590.) The development of the procedure, coupled with the availability of fresh calf bone, enabled isolation of substantially pure bovine osteogenic protein (BOP). BOP was characterized significantly; its ability to induce cartilage and ultimately endochondral bone growth in cat, rabbit, and rat were demonstrated and studied; it was shown to be able to induce the full developmental cascade of bone formation previously ascribed to unknown protein or proteins in heterogeneous bone extracts. This dose dependent and highly specific activity was present whether or not the protein was glycosylated (see U.S. Patent No. 4,968,958, filed 4/8/88 and Sampath et al., (1990) J. Biol. Chem. 265: pp. 13198-13205). Sequence data obtained from the bovine materials suggested probe designs which were used to isolate genes encoding osteogenic proteins from different species. Human and murine osteogenic protein counterparts have now been identified and characterized (see, for example, U.S. Pat. No. 5,011,691, Ozkaynak, et al., (1990) EMBO J 9:2085-2093, and Ozkaynak et al., (1991) Biochem. Biophys. Res. Commn. 179:116-123, and USSN 841,646, filed February 21, 1992, the disclosures of which are herein incorporated by reference.)

Sequence data from the bovine materials also suggested substantial homology with a number of proteins known in the art which were not known to play a role in bone formation. Bone formation assays

performed with these proteins showed that, when these proteins were implanted in a mammal in association with a suitable matrix, cartilage and endochondral bone formation was induced (see, for example, U.S. Patent  
5 No. 5,011,691.) One of these proteins is DPP, a *Drosophila* protein known to play a role in dorsal-ventral specification and required for the correct morphogenesis of the imaginal discs. Two other  
10 proteins are related sequences identified in *Xenopus* and mouse (Vgl and Vgr-1, respectively), thought to play a role in the control of growth and differentiation during embryogenesis. While DPP and Vgr-1 (or Vgr-1-like) transcripts have been identified in a variety of tissues (embryonic, neonatal and adult,  
15 Lyons et al., (1989) PNAS 86:4554-4 558, and see *infra*), Vgl transcripts, which are maternally inherited and spatially restricted to the vegetal endoderm, decline dramatically after gastrulation.

20 From these homologies a generic consensus sequence was derived which encompasses the active sequence required for inducing bone morphogenesis in a mammal when implanted in association with a matrix. The generic sequence has at least a conserved six  
25 cysteine skeleton (Generic Sequence 1, Seq. ID No. 1) or, optionally, a 7-cysteine skeleton (Generic Sequence 2, Seq. ID No. 2), which includes the conserved six cysteine skeleton defined by Generic Sequence 1, and an additional cysteine at residue 36,  
30 accomodating the additional cysteine residue identified in the OP2 proteins. Each "Xaa" in the generic sequences indicates that any one of the 20 naturally-occurring L-isomer,  $\alpha$ -amino acids or a derivative



thereof may be used at that position. Longer generic sequences which also are useful further comprise the following sequence at their N-termini:

5

Cys Xaa Xaa Xaa Xaa (Seq. ID No. 15)  
 1 5

Biosynthetic constructs designed from this  
 10 generic consensus sequence also have been shown to induce cartilage and/or endochondral bone formation (e.g., COP-1, COP-3, COP-4, COP-5, COP-7 and COP-16, described in U.S. Patent No. 5,011,691 and presented below in Table III.) Table II, set forth below,  
 15 compares the amino acid sequences of the active regions of native proteins that have been identified as morphogens, including human OP-1 (hOP-1, Seq. ID Nos. 5 and 16-17), mouse OP-1 (mOP-1, Seq. ID Nos. 6 and 18-19), human and mouse OP-2 (Seq. ID Nos. 7, 8, and 20-22), CBMP2A (Seq. ID No. 9), CBMP2B (Seq. ID No. 10),  
 20 DPP (from Drosophila, Seq. ID No. 11), Vgl, (from Xenopus, Seq. ID No. 12), Vgr-1 (from mouse, Seq. ID No. 13), and GDF-1 (Seq. ID No. 14.) In the table, three dots indicates that the amino acid in that  
 25 position is the same as the amino acid in hOP-1. Three dashes indicates that no amino acid is present in that position, and are included for purposes of illustrating homologies. For example, amino acid residue 60 of CBMP-2A and CBMP-2B is "missing". Of course, both  
 30 these amino acid sequences in this region comprise Asn-Ser (residues 58, 59), with CBMP-2A then comprising Lys and Ile, whereas CBMP-2B comprises Ser and Ile.

TABLE II

	hOP-1	Cys	Lys	Lys	His	Glu	Leu	Tyr	Val	
	mOP-1	...	...	...	...	...	...	...	...	
5	hOP-2	...	Arg	Arg	...	...	...	...	...	
	mOP-2	...	Arg	Arg	...	...	...	...	...	
	DPP	...	Arg	Arg	...	Ser	...	...	...	
	Vgl	...	...	Lys	Arg	His	...	...	...	
	Vgr-1	...	...	...	...	Gly	...	...	...	
10	CBMP-2A	...	...	Arg	...	Pro	...	...	...	
	CBMP-2B	...	Arg	Arg	...	Ser	...	...	...	
	GDF-1	...	Arg	Ala	Arg	Arg	...	...	...	
		1				5				
15	hOP-1	Ser	Phe	Arg	Asp	Leu	Gly	Trp	Gln	Asp
	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	...	Gln	...	...	...	...	Leu	...
	mOP-2	Ser	...	...	...	...	...	...	Leu	...
20	DPP	Asp	...	Ser	...	Val	...	...	Asp	...
	Vgl	Glu	...	Lys	...	Val	...	...	...	Asn
	Vgr-1	...	...	Gln	...	Val	...	...	...	...
	CBMP-2A	Asp	...	Ser	...	Val	...	...	Asn	...
	CBMP-2B	Asp	...	Ser	...	Val	...	...	Asn	...
25	GDF-1	...	...	...	Glu	Val	...	...	His	Arg
			10					15		
	hOP-1	Trp	Ile	Ile	Ala	Pro	Glu	Gly	Tyr	Ala
	mOP-1	...	...	...	...	...	...	...	...	...
30	hOP-2	...	Val	...	...	...	Gln	...	...	Ser
	mOP-2	...	Val	...	...	...	Gln	...	...	Ser
	DPP	...	...	Val	...	...	Leu	...	...	Asp
	Vgl	...	Val	...	...	...	Gln	...	...	Met
	Vgr-1	...	...	...	...	...	Lys	...	...	...

	CBMP-2A	...	...	Val	...	...	Pro	...	...	His
	CBMP-2B	...	...	Val	...	...	Pro	...	...	Gln
	GDF-1	...	Val	...	...	...	Arg	...	Phe	Leu
				20					25	
5										
	hOP-1	Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ala
	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	...	...	...	...	...	...	...	Ser
10	mOP-2	...	...	...	...	...	...	...	...	...
	DPP	...	...	...	...	His	...	Lys	...	Pro
	Vgl	...	Asn	...	...	Tyr	...	...	...	Pro
	Vgr-1	...	Asn	...	...	Asp	...	...	...	Ser
	CBMP-2A	...	Phe	...	...	His	...	Glu	...	Pro
15	CBMP-2B	...	Phe	...	...	His	...	Asp	...	Pro
	GDF-1	...	Asn	...	...	Gln	...	Gln	...	...
				30						35
	hOP-1	Phe	Pro	Leu	Asn	Ser	Tyr	Met	Asn	Ala
20	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	...	...	Asp	...	Cys	...	...	...
	mOP-2	...	...	...	Asp	...	Cys	...	...	...
	DPP	...	...	...	Ala	Asp	His	Phe	...	Ser
	Vgl	Tyr	...	...	Thr	Glu	Ile	Leu	...	Gly
25	Vgr-1	...	...	...	...	Ala	His	...	...	...
	CBMP-2A	...	...	...	Ala	Asp	His	Leu	...	Ser
	CBMP-2B	...	...	...	Ala	Asp	His	Leu	...	Ser
	GDF-1	Leu	...	Val	Ala	Leu	Ser	Gly	Ser**	...
					40					
30										
	hOP-1	Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu
	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	...	...	...	...	Leu	...	Ser	...
	mOP-2	...	...	...	...	...	Leu	...	Ser	...
35	DPP	...	...	...	...	Val	...	...	...	...

[illegible]

	hOP-2	...	Ser	...	Thr	...	...	...	...	Tyr
	mOP-2	...	Ser	...	Thr	...	...	...	...	Tyr
	Vgl	Met	Ser	Pro	...	...	Met	...	Phe	Tyr
	Vgr-1	Val	...	...	...	...	...	...	...	...
5	DPP	...	Asp	Ser	Val	Ala	Met	...	...	Leu
	CBMP-2A	...	Ser	...	...	...	Met	...	...	Leu
	CBMP-2B	...	Ser	...	...	...	Met	...	...	Leu
	GDF-1	...	Ser	Pro	...	...	...	...	Phe	...
					75					80
10	hOP-1	Asp	Asp	Ser	Ser	Asn	Val	Ile	Leu	Lys
	mOP-1	...	...	...	...	...	...	...	...	...
	hOP-2	...	Ser	...	Asn	...	...	...	...	Arg
	mOP-2	...	Ser	...	Asn	...	...	...	...	Arg
	DPP	Asn	...	Gln	...	Thr	...	Val	...	...
15	Vgl	...	Asn	Asn	Asp	...	...	Val	...	Arg
	Vgr-1	...	...	Asn	...	...	...	...	...	...
	CBMP-2A	...	Glu	Asn	Glu	Lys	...	Val	...	...
	CBMP-2B	...	Glu	Tyr	Asp	Lys	...	Val	...	...
	GDF-1	...	Asn	...	Asp	...	...	Val	...	Arg
20						85				
	hOP-1	Lys	Tyr	Arg	Asn	Met	Val	Val	Arg	
	mOP-1	...	...	...	...	...	...	...	...	
25	hOP-2	...	His	...	...	...	...	...	Lys	
	mOP-2	...	His	...	...	...	...	...	Lys	
	DPP	Asn	...	Gln	Glu	...	Thr	...	Val	
	Vgl	His	...	Glu	...	...	Ala	...	Asp	
	Vgr-1	...	...	...	...	...	...	...	...	
30	CBMP-2A	Asn	...	Gln	Asp	...	...	...	Glu	
	CBMP-2B	Asn	...	Gln	Glu	...	...	...	Glu	
	GDF-1	Gln	...	Glu	Asp	...	...	...	Asp	
		90							95	

	hOP-1	Ala	Cys	Gly	Cys	His
	mOP-1	...	...	...	...	...
	hOP-2	...	...	...	...	...
	mOP-2	...	...	...	...	...
5	DPP	Gly	...	...	...	Arg
	Vgl	Glu	...	...	...	Arg
	Vgr-1	...	...	...	...	...
	CBMP-2A	Gly	...	...	...	Arg
	CBMP-2B	Gly	...	...	...	Arg
10	GDF-1	Glu	...	...	...	Arg

100

\*\*Between residues 43 and 44 of GDF-1 lies the amino acid sequence Gly-Gly-Pro-Pro.

15 Table III, set forth below, compares the amino acid sequence data for six related biosynthetic constructs designated COPs 1, 3, 4, 5, 7, and 16. These sequences also are presented in U.S. Pat. No. 5,011,691. As with Table II, the dots mean that in that position there is an identical amino acid to that of COP-1, and dashes mean that the COP-1 amino acid is missing at that position.

25

TABLE III

	COP-1	Leu	Tyr	Val	Asp	Phe	Gln	Arg	Asp	Val
	COP-3	...	...	...	...	...	...	...	...	...
	COP-4	...	...	...	...	...	Ser	---	...	...
30	COP-5	...	...	...	...	...	Ser	---	...	...
	COP-7	...	...	...	...	...	Ser	---	...	...
	COP-16	...	...	...	...	...	Ser	---	...	...
		1				5				

5	COP-1	Gly	Trp	Asp	Asp	Trp	Ile	Ile	Ala
	COP-3	...	...	...	...	...	...	Val	...
	COP-4	...	...	...	...	...	...	Val	...
	COP-5	...	...	...	...	...	...	Val	...
	COP-7	...	...	Asn	...	...	...	Val	...
	COP-16	...	...	Asn	...	...	...	Val	...
		10					15		
10	COP-1	Pro	Val	Asp	Phe	Asp	Ala	Tyr	Tyr
	COP-3	...	Pro	Gly	Tyr	Gln	...	Phe	...
	COP-4	...	Pro	Gly	Tyr	Gln	...	Phe	...
	COP-5	...	Pro	Gly	Tyr	Gln	...	Phe	...
	COP-7	...	Pro	Gly	Tyr	His	...	Phe	...
	COP-16	...	Pro	Gly	Tyr	Gln	...	Phe	...
15				20					25
20	COP-1	Cys	Ser	Gly	Ala	Cys	Gln	Phe	Pro
	COP-3	...	...	...	...	...	...	...	...
	COP-4	...	...	...	...	...	...	...	...
	COP-5	...	His	...	Glu	...	Pro	...	...
	COP-7	...	His	...	Glu	...	Pro	...	...
	COP-16	...	His	...	Glu	...	Pro	...	...
				30					
25									
30	COP-1	Ser	Ala	Asp	His	Phe	Asn	Ser	Thr
	COP-3	...	...	...	...	...	...	...	...
	COP-4	...	...	...	...	...	...	...	...
	COP-5	Leu	...	...	...	...	...	...	...
	COP-7	Leu	...	...	...	Leu	...	...	...
	COP-16	Leu	...	...	...	...	...	...	...
			35					40	

		Asn	His	Ala	Val	Val	Gln	Thr	Leu	Val
5	COP-1	...	...	...	...	...	...	...	...	...
	COP-3	...	...	...	...	...	...	...	...	...
	COP-4	...	...	...	...	...	...	...	...	...
	COP-5	...	...	...	...	...	...	...	...	...
	COP-7	...	...	...	...	...	...	...	...	...
	COP-16	...	...	...	...	...	...	...	...	...
				45						50
10	COP-1	Asn	Asn	Met	Asn	Pro	Gly	Lys	Val	
	COP-3	...	...	...	...	...	...	...	...	
	COP-4	...	...	...	...	...	...	...	...	
	COP-5	...	Ser	Val	...	Ser	Lys	Ile	---	
	COP-7	...	Ser	Val	...	Ser	Lys	Ile	---	
	COP-16	...	Ser	Val	...	Ser	Lys	Ile	---	
15					55					
20	COP-1	Pro	Lys	Pro	Cys	Cys	Val	Pro	Thr	
	COP-3	...	...	...	...	...	...	...	...	
	COP-4	...	...	...	...	...	...	...	...	
	COP-5	...	...	Ala	...	...	...	...	...	
	COP-7	...	...	Ala	...	...	...	...	...	
	COP-16	...	...	Ala	...	...	...	...	...	
25			60					65		
30	COP-1	Glu	Leu	Ser	Ala	Ile	Ser	Met	Leu	
	COP-3	...	...	...	...	...	...	...	...	
	COP-4	...	...	...	...	...	...	...	...	
	COP-5	...	...	...	...	...	...	...	...	
	COP-7	...	...	...	...	...	...	...	...	
	COP-16	...	...	...	...	...	...	...	...	
35				70						



	COP-1	Tyr	Leu	Asp	Glu	Asn	Ser	Thr	Val
	COP-3	...	...	...	...	...	Glu	Lys	...
	COP-4	...	...	...	...	...	Glu	Lys	...
5	COP-5	...	...	...	...	...	Glu	Lys	...
	COP-7	...	...	...	...	...	Glu	Lys	...
	COP-16	...	...	...	...	...	Glu	Lys	...
		75					80		
10									
	COP-1	Val	Leu	Lys	Asn	Tyr	Gln	Glu	Met
	COP-3	...	...	...	...	...	...	...	...
	COP-4	...	...	...	...	...	...	...	...
	COP-5	...	...	...	...	...	...	...	...
15	COP-7	...	...	...	...	...	...	...	...
	COP-16	...	...	...	...	...	...	...	...
				85					90
20									
	COP-1	Thr	Val	Val	Gly	Cys	Gly	Cys	Arg
	COP-3	Val	...	Glu	...	...	...	...	...
	COP-4	Val	...	Glu	...	...	...	...	...
	COP-5	Val	...	Glu	...	...	...	...	...
	COP-7	Val	...	Glu	...	...	...	...	...
25	COP-16	Val	...	Glu	...	...	...	...	...
					95				

30           As is apparent from the foregoing amino acid  
sequence comparisons, significant amino acid changes  
can be made within the generic sequences while  
retaining the morphogenic activity. For example, while  
the GDF-1 protein sequence depicted in Table II shares  
35 only about 50% amino acid identity with the hOP1

sequence described therein, the GDF-1 sequence shares greater than 70% amino acid sequence homology with the hOP1 sequence, where homology is defined by allowed conservative amino acid changes within the sequence as  
5 defined by Dayoff, et al., Atlas of Protein Sequence and Structure vol.5, supp.3, pp.345-362, (M.O. Dayoff, ed., Nat'l BioMed. Res. Fd'n, Washington D.C. 1979.)

It now has been discovered that the family of  
10 proteins described by these sequences also is capable of initiating and maintaining the tissue-specific developmental cascade in tissues other than bone and cartilage. When combined with naive progenitor cells as disclosed herein, these proteins, termed morphogens,  
15 are capable of inducing the proliferation and differentiation of the progenitor cells. In the presence of appropriate tissue-specific signals to direct the differentiation of these cells, and a morphogenically permissive environment, these  
20 morphogens are capable of reproducing the cascade of cellular and molecular events that occur during embryonic development to yield functional tissue.

A key to these developments was the creation  
25 of a mammalian tissue model system, namely a model system for endochondral bone formation, and investigation of the cascade of events important for bone tissue morphogenesis. Work on this system has enabled discovery not only of bone inductive  
30 morphogens, but also of tissue inductive morphogens and their activities. The methods used to develop the bone model system, now well known in the art, along with the proteins of this invention, can be used to create model systems for other tissues, such as liver (see infra).

Using the model system for endochondral bone formation, it also has been discovered that the local environment in which the morphogenic material is placed is important for tissue morphogenesis. As used herein, "local environment" is understood to include the tissue structural matrix and the environment surrounding the tissue. For example, in addition to needing an appropriate anchoring substratum for their proliferation, the morphogen-stimulated cells need signals to direct the tissue-specificity of their differentiation. These signals vary for the different tissues and may include cell surface markers. In addition, vascularization of new tissue requires a local environment which supports vascularization. Using the bone model system as an example, it is known that, under standard assay conditions, implanting osteoinductive morphogens into loose mesenchyme in the absence of a tissue-specifying matrix generally does not result in endochondral bone formation unless very high concentrations of the protein are implanted. By contrast, implanting relatively low concentrations of the morphogen in association with a suitably modified bone-derived matrix results in the formation of fully functional endochondral bone (see, for example, Sampath et al. (1981) PNAS 78:7599-7 603 and U.S. Patent No. 4,975,526). In addition, a synthetic matrix comprised of a structural polymer such as tissue-specific collagen and tissue-specific cell attachment factors such as tissue-specific glycosylaminoglycans, will allow endochondral bone formation (see, for example, PCT publication US91/03603, published December 12, 1991 (WO 91/18558), incorporated herein by reference). Finally, if the morphogen and a suitable bone or cartilage-specific matrix (e.g., comprising Type I cartilage) are implanted together in loose mesenchyme, cartilage and endochondral bone formation will result, including the formation of bone marrow and

a vascular system. However, if the same composition is provided to a nonvascular environment, such as to cultured cells in vitro or at an cartilage-specific locus, tissue development does not continue beyond cartilage formation (see *infra*). Similarly, a morphogenic composition containing a cartilage-specific matrix composed of Type 2 collagen is expected to induce formation of non-cartilage tissue in vivo (e.g., hyaline). However, if the composition is provided to a vascular-supporting environment, such as loose mesenchyme, the composition is capable of inducing the differentiation of proliferating progenitor cells into chondrocytes and osteoblasts, resulting in bone formation.

It also has been discovered that tissue morphogenesis requires a morphogenically permissive environment. Clearly, in fully-functioning healthy tissue that is not composed of a permanently renewing cell population, there must exist signals to prevent continued tissue growth. Thus, it is postulated that there exists a control mechanism, such as a feedback control mechanism, which regulates the control of cell growth and differentiation. In fact, it is known that both TGF- $\beta$ , and MIS are capable of inhibiting cell growth when present at appropriate concentrations. In addition, using the bone model system it can be shown that osteogenic devices comprising a bone-derived carrier (matrix) that has been demineralized and guanidine-extracted to substantially remove the noncollagenous proteins does allow endochondral bone formation when implanted in association with an

osteoinductive morphogen. If, however, the bone-derived carrier is not demineralized but rather is washed only in low salt, for example, induction of endochondral bone formation is inhibited, suggesting  
5 the presence of one or more inhibiting factors within the carrier.

Another key to these developments was determination of the broad distribution of these  
10 morphogens in developing and adult tissue. For example, DPP is expressed in both embryonic and developing *Drosophila* tissue. Vgl has been identified in *Xenopus* embryonic tissue. Vgr-1 transcripts have been identified in a variety of murine tissues,  
15 including embryonic and developing brain, lung, liver, kidney and calvaria (dermal bone) tissue. Recently, Vgr-1 transcripts also have been identified in adult murine lung, kidney, heart, and brain tissue, with especially high abundance in the lung (see infra).

20 OP-1 and the CBMP2 proteins, both first identified as bone morphogens, have been identified in mouse and human placenta, hippocampus, calvaria and osteosarcoma tissue as determined by identification of  
25 OP-1 and CBMP2-specific sequences in cDNA libraries constructed from these tissues (see Ozkaynak, et al., (1990) EMBO J 9:2085-2093, and Ozkaynak et al., (1991) Biochem. Biophys. Res. Commun. 179:116-123). Additionally, the OP-1 protein is present in a variety  
30 of embryonic and developing tissues including kidney, liver, heart, adrenal tissue and brain as determined by Western blot analysis and immunolocalization (see infra). OP-1-specific transcripts also have been identified in both embryonic and developing tissues,  
35 most abundantly in developing kidney, bladder and brain

(see infra). OP-1 also has been identified as a mesoderm inducing factor present during embryogenesis (see infra). Moreover, OP-1 has been shown to be associated with in satellite muscle cells and  
5 associated with pluripotential stem cells in bone marrow following damage to adult murine endochondral bone, indicating its morphogenic role in tissue repair and regeneration. In addition, the recently identified protein GDF-1 (see Table II) has been identified in  
10 neural tissue (Lee, (1991) PNAS 88 4250-4254).

### Exemplification

#### IDENTIFICATION AND ISOLATION OF MORPHOGENS

15

Among the proteins useful in this invention are proteins originally identified as bone inductive proteins, such as the OP-1, OP-2 and the CBMP proteins, as well as amino acid sequence-related proteins such as  
20 DPP (from *Drosophila*), Vgl (from *Xenopus*) and Vgr-1 (from mouse, see Table II and Sequence Listing). The members of this family, which include particular members of the TGF- $\beta$  super family of structurally related proteins, share substantial amino acid sequence  
25 homology in their C-terminal regions. The OP-2 proteins have an extra cysteine residue in this region (position 41 of Seq. ID Nos. 7 and 8), in addition to the conserved cysteine skeleton in common with the other proteins in this family. The proteins are  
30 inactive when reduced, but are active as oxidized homodimeric species as well as when oxidized in combination with other morphogens.

Accordingly, the morphogens of this invention  
35 can be described by either of the following two species

of generic amino acid sequences: Generic Sequence 1 or Generic Sequence 2, (Seq. ID Nos. 1 and 2), where each Xaa indicates one of the 20 naturally-occurring L-isomer,  $\alpha$ -amino acids or a derivative thereof.

5 Particularly useful sequences that fall within this family of proteins include the 96-102 C-terminal residues of Vgl, Vgr-1, DPP, OP-1, OP-2, CBMP-2A, CBMP-2B, and GDF-1, as well as their intact mature amino acid sequences. In addition, biosynthetic

10 constructs designed from the generic sequences, such as COP-1, COP-3-5, COP-7, and COP-16 also are useful (see, for example, U.S. Pat. No. 5,011,691.)

Generic sequences showing preferred amino

15 acids compiled from sequences identified to date and useful as morphogens (e.g., Tables II and III) are described by Generic Sequence 3 (Seq. ID No. 3) and Generic Sequence 4 (Seq. ID No. 4). Note that these generic sequences have a 7 or 8-cysteine skeleton where

20 inter- or intramolecular disulfide bonds can form, and contain certain critical amino acids which influence the tertiary structure of the proteins. It is also contemplated that the differing N-termini of the naturally occurring proteins provide a tissue-specific

25 or other, important modulating activity of these proteins.

Given the foregoing amino acid and DNA sequence information, the level of skill in the art,

30 and the disclosures of U.S. Patent Nos. 4,968,590 and 5,011,691, PCT application US 89/01469, published October 19, 1989 (WO89/09788), and Ozkaynak, et al., (1990) EMBO J 9:2085-2093, and Ozkaynak et al., (1991) Biochem. Biophys. Res. Commun. 179:116-123 the

35 disclosures of which are incorporated herein by reference, various DNAs can be constructed which encode

at least the active region of a morphogen of this invention, and various analogs thereof (including allelic variants and those containing genetically engineered mutations), as well as fusion proteins, truncated forms of the mature proteins, deletion and insertion mutants, and similar constructs. Moreover, DNA hybridization probes can be constructed from fragments of the genes encoding any of these proteins, including sequences encoding the active regions or the pro regions of the proteins (see *infra*), or designed de novo from the generic sequence. These probes then can be used to screen different genomic and cDNA libraries to identify additional morphogenic proteins from different tissues.

15

The DNAs can be produced by those skilled in the art using well known DNA manipulation techniques involving genomic and cDNA isolation, construction of synthetic DNA from synthesized oligonucleotides, and cassette mutagenesis techniques. 15-100mer oligonucleotides may be synthesized on a Biosearch DNA Model 8600 Synthesizer, and purified by polyacrylamide gel electrophoresis (PAGE) in Tris-Borate-EDTA buffer. The DNA then may be electroeluted from the gel. Overlapping oligomers may be phosphorylated by T4 polynucleotide kinase and ligated into larger blocks which also may be purified by PAGE.

The DNA from appropriately identified clones then can be isolated, subcloned (preferably into an expression vector), and sequenced. Plasmids containing sequences of interest then can be transfected into an appropriate host cell for expression of the morphogen and further characterization. The host may be a



procaryotic or eucaryotic cell since the former's inability to glycosylate protein will not destroy the protein's morphogenic activity. Useful host cells include E. coli, Saccharomyces, the insect/baculovirus cell system, myeloma cells, and various other mammalian cells. The vectors additionally may encode various sequences to promote correct expression of the recombinant protein, including transcription promoter and termination sequences, enhancer sequences, preferred ribosome binding site sequences, preferred mRNA leader sequences, preferred signal sequences for protein secretion, and the like.

The DNA sequence encoding the gene of interest also may be manipulated to remove potentially inhibiting sequences or to minimize unwanted secondary and tertiary structure formation. The recombinant morphogen also may be expressed as a fusion protein. After being translated, the protein may be purified from the cells themselves or recovered from the culture medium. All biologically active protein forms comprise dimeric species joined by disulfide bonds or otherwise associated, produced by refolding and oxidizing one or more of the various recombinant polypeptide chains within an appropriate eucaryotic cell or in vitro after expression of individual subunits. A detailed description of morphogens expressed from recombinant DNA in E. coli and in numerous different mammalian cells is disclosed in PCT publication US90/05903, published May 2, 1991 (WO91/05802) and U.S. Serial No. 841,646 filed February 21, 1992, the disclosures of which are hereby incorporated by reference.

Alternatively, morphogenic polypeptide chains can be synthesized chemically using conventional peptide synthesis techniques well known to those having

ordinary skill in the art. For example, the proteins may be synthesized intact or in parts on a Biosearch solid phase peptide synthesizer, using standard operating procedures. Completed chains then are  
5 deprotected and purified by HPLC (high pressure liquid chromatography). If the protein is synthesized in parts, the parts may be peptide bonded using standard methodologies to form the intact protein. In general, the manner in which the morphogens are made can be  
10 conventional and does not form a part of this invention.

#### MORPHOGEN DISTRIBUTION

15 The generic function of the morphogens of this invention throughout the life of the organism can be evidenced by their expression in a variety of disparate mammalian tissues. Determination of the tissue distribution of morphogens also may be used to identify  
20 different morphogens expressed in a given tissue, as well as to identify new, related morphogens. The proteins (or their mRNA transcripts) are readily identified in different tissues using standard methodologies and minor modifications thereof in  
25 tissues where expression may be low. For example, protein distribution may be determined using standard Western blot analysis or immunofluorescent techniques, and antibodies specific to the morphogen or morphogens of interest. Similarly, the distribution of morphogen  
30 transcripts may be determined using standard Northern hybridization protocols and transcript-specific probes.

Any probe capable of hybridizing specifically to a transcript, and distinguishing the transcript of  
35 interest from other, related transcripts may be used.

Because the morphogens of this invention share such high sequence homology in their active, C-terminal domains, the tissue distribution of a specific morphogen transcript may best be determined using a probe specific for the pro region of the immature protein and/or the N-terminal region of the mature protein. Another useful sequence is the 3' non-coding region flanking and immediately following the stop codon. These portions of the sequence vary substantially among the morphogens of this invention, and accordingly, are specific for each protein. For example, a particularly useful Vgr-1-specific probe sequence is the PvuII-SacI fragment, a 265 bp fragment encoding both a portion of the untranslated pro region and the N-terminus of the mature sequence (see Lyons et al. (1989) PNAS 86:4554-4558 for a description of the cDNA sequence). Similarly, particularly useful mOP-1-specific probe sequences are the BstXI-BglI fragment, a 0.68 Kb sequence that covers approximately two-thirds of the mOP-1 pro region; a StuI-StuI fragment, a 0.2 Kb sequence immediately upstream of the 7-cysteine domain; and the Earl-PstI fragment, an 0.3 Kb fragment containing a portion of the 3'untranslated sequence (See Seq. ID No. 18, where the pro region is defined essentially by residues 30-291.) Similar approaches may be used, for example, with hOP1 (Seq. ID No. 16) or human or mouse OP2 (Seq. ID Nos. 20 and 22.)

Using these morphogen-specific probes, which may be synthetically engineered or obtained from cloned sequences, morphogen transcripts can be identified in mammalian tissue, using standard methodologies well known to those having ordinary skill in the art. Briefly, total RNA is prepared from various adult murine tissues (e.g., liver, kidney, testis, heart,

brain, thymus and stomach) by a standard methodology such as by the method of Chomczyaski et al. ((1987) Anal. Biochem 162:156-159) and described below. Poly (A)+ RNA is prepared by using oligo (dT)-cellulose chromatography (e.g., Type 7, from Pharmacia LKB Biotechnology, Inc.). Poly (A)+ RNA (generally 15 µg) from each tissue is fractionated on a 1% agarose/formaldehyde gel and transferred onto a Nytran membrane (Schleicher & Schuell). Following the transfer, the membrane is baked at 80°C and the RNA is cross-linked under UV light (generally 30 seconds at 1 mW/cm<sup>2</sup>). Prior to hybridization, the appropriate probe (e.g., the PvuII-SacI Vgr-1 fragment) is denatured by heating. The hybridization is carried out in a lucite cylinder rotating in a roller bottle apparatus at approximately 1 rev/min for approximately 15 hours at 37°C using a hybridization mix of 40% formamide, 5 x Denhardts, 5 x SSPE, and 0.1% SDS. Following hybridization, the non-specific counts are washed off the filters in 0.1 x SSPE, 0.1% SDS at 50°C. Northern blots performed using Vgr-1 probes specific to the variable N terminus of the mature sequence indicate that the Vgr-1 message is approximately 3.5 Kb.

Figure 1 is a photomicrograph representing a Northern blot analysis probing a number of adult murine tissues with the Vgr-1 specific probes: liver, kidney, testis, heart, brain, thymus and stomach, represented in lanes 3-10, respectively. Lanes 1 and 12 are size standards and lanes 2 and 11 are blank. Among the tissues tested, Vgr-1 appears to be expressed most abundantly in adult lung, and to a lesser extent in adult kidney, heart and brain. These results confirm and expand on earlier studies identifying Vgr-1 and Vgr-1-like transcripts in several embryonic and adult

murine tissue (Lyons et al. (1989) PNAS 86:4554-4558),  
as well as studies identifying OP-1 and CBMP2 in  
various human cDNA libraries (e.g., placenta,  
hippocampus, calvaria, and osteosarcoma, see Ozkaynak  
5 et al., (1990) EMBO 9:2085-2093).

Using the same general probing methodology,  
mOP-1 transcripts also have been identified in a  
variety of murine tissues, including embryo and various  
10 developing tissues, as can be seen in Figures 2 and 3.  
Details of the probing methodology are disclosed in  
Ozkaynak, et al., (1991) Biochem. Biophys. Res. Commn.  
179:116-123, the disclosure of which is incorporated  
herein. The Northern blots represented in Figure 2  
15 probed RNA prepared from developing brain, spleen,  
lung, kidney (and adrenal gland), heart, and liver in  
13 day post natal mice (panel A) or 5 week old mice  
(panel B). The OP-1 specific probe was a probe  
containing the 3' untranslated sequences described  
20 supra (0.34 Kb EarI-Pst I fragment). As a control for  
RNA recovery, EF-Tu (translational elongation factor)  
mRNA expression also was measured (EF-Tu expression is  
assumed to be relatively uniform in most tissues).

25 The arrowheads indicate the OP1-specific messages  
observed in the various tissues. As can be seen in  
Fig. 2, OP-1 expression levels vary significantly in  
the spleen, lung, kidney and adrenal tissues, while the  
EF-Tu mRNA levels are constant. Uniformly lower levels  
30 of EF-Tu mRNA levels were found in the heart, brain and  
liver. As can be seen from the photomicrograph, the  
highest levels of OP-1 mRNA appear to be in kidney and  
adrenal tissue, followed by the brain. By contrast,  
heart and liver did not give a detectable signal. Not

shown are additional analyses performed on bladder tissue, which shows significant OP-1 mRNA expression, at levels close to those in kidney/adrenal tissue. The Northern blots also indicate that, like GDF-1, OP-1 mRNA expression may be bicistronic in different tissues. Four transcripts can be seen: 4 Kb, 2.4 Kb, 2.2 Kb, and 1.8 Kb transcripts can be identified in the different tissues, and cross probing with OP-1 specific probes from the proregion and N-terminal sequences of the gene indicate that these transcripts are OP-1 specific.

A side by side comparison of OP-1 and Vgr-1 in Figure 3 shows that the probes distinguish between the morphogens Vgr-1 and OP-1 transcripts in the different tissues, and also highlights the multiple transcription of OP-1 in different tissues. Specifically, Fig. 3 compares the expression of OP-1 (Panels B and D), Vgr-1 (Panel C) and EF-Tu (Panel A) (control) mRNA in 17 day embryos (lane 1) and 3 day post-natal mice (lane 2). The same filter was used for sequential hybridizations with labeled DNA probes specific for OP-1 (Panels B and D), Vgr-1 (Panel C), and EF-Tu (Panel A). Panel A: the EF-Tu specific probe (control) was the 0.4 Kb HindIII-SacI fragment (part of the protein coding region), the SacI site used belonged to the vector; Panel B: the OP-1 specific probe was the 0.68 Kb BstXI-BglI fragment containing pro region sequences; Panel D; the OP-1 specific probe was the 0.34 Kb EarI-PstI fragment containing the 3' untranslated sequence; Panel C: the Vgr-1 specific probe was the 0.26 Kb PvuII-SacI fragment used in the Vgr-1 blots described above.

The 1.8-2.5 Kb OP-1 mRNA appears approximately two times higher in three day post natal mice than in 17 day embryos, perhaps reflecting phases in bone and/or kidney development. In addition, of the four messages  
5 found in brain, the 2.2 Kb transcript appears most abundant, whereas in lung and spleen the 1.8 Kb message predominates. Finally, careful separation of the renal and adrenal tissue in five week old mice reveals that the 2.2 Kb transcripts were derived from renal tissue  
10 and the 4 Kb mRNA is more prominent in adrenal tissue (see Figure 2).

Similarly, using the same general probing methodology, BMP3 and CBMP2B transcripts recently have  
15 been identified in abundance in lung tissue.

Morphogen distribution in embryonic tissue can be determined using five or six-day old mouse embryos and standard immunofluorescence techniques in concert  
20 with morphogen-specific antisera. For example, rabbit anti-OP-1 antisera is readily obtained using any of a number of standard antibody protocols well known to those having ordinary skill in the art. The antibodies then are fluorescently labelled using standard  
25 procedures. A five or six-day old mouse embryo then is thin-sectioned and the various developing tissues probed with the labelled antibody, again following standard protocols. Using this technique, OP-1 protein has been detected in developing brain and heart.

30

This method also may be used to identify morphogens in adult tissues undergoing repair. For example, a fracture site can be induced in a rat long bone such as the femur. The fracture then is allowed  
35 to heal for 2 or 3 days. The animal then is sacrificed

and the fractured site sectioned and probed for the presence of the morphogen e.g., OP-1, with fluorescently labelled rabbit anti-OP-1 antisera using standard immunolocalization methodology. This  
5 technique identifies OP-1 in muscle satellite cells, the progenitor cells for the development of muscle, cartilage and endochondral bone. In addition, OP-1 is detected with potential pluripotential stem cells in the bone marrow, indicating its morphogenic role in  
10 tissue repair and regeneration.

OP-1 protein also has been identified in rat brain using standard immunofluorescence staining technique. Specifically, adult rat brain (2-3 months old) and  
15 spinal cord is frozen and sectioned. Anti-OP-1, raised in rabbits and purified on an OP-1 affinity column prepared using standard methodologies, was added to the sections under standard conditions for specific binding. Goat anti-rabbit IgG, labelled with  
20 fluorescence, then was used to visualize OP-1 antibody binding to tissue sections.

As can be seen in FIG 4A and 4B, immunofluorescence staining demonstrates the presence of OP-1 in adult rat  
25 central nervous system (CNS.) Similar and extensive staining is seen in both the brain (4A) and spinal cord (4B). OP-1 appears to be predominantly localized to the extracellular matrix of the grey matter, present in all areas except the neuronal cell bodies. In white  
30 matter, staining appears to be confined to astrocytes. A similar staining pattern also was seen in newborn rat (10 day old) brain sections.

#### CELL DIFFERENTIATION



The ability of morphogens of this invention to induce cell differentiation can be determined by culturing early mesenchymal cells in the presence of the morphogen and then studying the histology of the cultured cells by staining with toluidine blue. For example, it is known that rat mesenchymal cells destined to become mandibular bone, when separated from the overlying epithelial cells at stage 11 and cultured in vitro under standard tissue culture conditions, will not continue to differentiate. However, if these same cells are left in contact with the overlying endoderm for an additional day, at which time they become stage 12 cells, they will continue to differentiate on their own in vitro to form chondrocytes. Further differentiation into osteoblasts and, ultimately, mandibular bone, requires an appropriate local environment, e.g., a vascularized environment.

It has now been discovered that stage 11 mesenchymal cells, cultured in vitro in the presence of a morphogen, e.g., OP-1, continue to differentiate in vitro to form chondrocytes. These stage 11 cells also continue to differentiate in vitro if they are cultured with the cell products harvested from the overlying endodermal cells. Moreover, OP-1 can be identified in the medium conditioned by endodermal cells either by Western blot or immunofluorescence. This experiment may be performed with other morphogens and with different mesenchymal cells to assess the cell differentiation capability of different morphogens, as well as their distribution in different developing tissues.

As another example of morphogen-induced cell differentiation, the effect of OP-1 on the

differentiation of neuronal cells has been tested in culture. Specifically, the effect of OP-1 on the NG108-15 neuroblastoma x glioma hybrid clonal cell line has been assessed. The cell line shows a fibroblastic-  
5 type morphology in culture. The cell line can be induced to differentiate chemically using 0.5 mM butyrate, 1% DMSO or 500 mM Forskolin, inducing the expression of virtually all important neuronal properties of cultured primary neurons. However,  
10 chemical induction of these cells also induces cessation of cell division.

In the present experiment NG108-15 cells were subcultured on poly-L-lysine coated 6 well plates.  
15 Each well contained 40-50,000 cells in 2.5 ml of chemically defined medium. On the third day 2.5  $\mu$ l of OP-1 in 60% ethanol containing 0.025% trifluoroacetic was added to each well. OP-1 concentrations of 0, 1, 10, 40 and 100 ng/ml were tested. The media was  
20 changed daily with new aliquots of OP-1. After four days with 40 and 100 ng OP-1/ml concentrations, OP-1 induced differentiation of the NG108-15 cells. Figure 5 shows the morphological changes that occur. The OP-1 induces clumping and rounding of the cells and  
25 the production of neurite outgrowths (processes). Compare FIG 5A (naive NG108-15 cells) with FIG 5B, showing the effects of OPI-treated cells. Thus the OP-1 can induce the cells to differentiate into a neuronal cell morphology. Some of the outgrowths  
30 appear to join in a synaptic-type junction. This effect was not seen in cells incubated with TGF-B1 at concentrations of 1 to 100 ng/ml.

The neuroprotective effects of OP-1 were  
35 demonstrated by comparison with chemical

differentiation agents on the NG108-15 cells. 50,000 cells were plated on 6 well plates and treated with butyrate, DMSO, Forskolin or OP-1 for four days. Cell counts demonstrated that in the cultures containing the chemical agents the differentiation was accompanied by a cessation of cell division. In contrast, the cells induced to differentiate by OP-1 continued to divide, as determined by  $H^3$ -thymidine uptake. The data suggest that OP-1 is capable of maintaining the stability of the cells in culture after differentiation.

As yet another, related example, the ability of the morphogens of this invention to induce the "redifferentiation" of transformed cells also has been assessed. Specifically, the effect of OP-1 on human EC cells (embryo carcinoma cells, NTERA-2 CL.D1) is disclosed herein. In the absence of an external stimulant these cells can be maintained as undifferentiated stem cells, and can be induced to grow in serum free media (SFM). In the absence of morphogen treatment the cells proliferate rampantly and are anchorage-independent. The effect of morphogen treatment is seen in Figs. 6A-D. Figs 6A and 6B show 4 days of growth in SFM in the presence of OP-1 (25ng/ml, 6A) or the absence of morphogen (6B). Figs. 6C and 6D are 5 days growth in the presence of 10ng/ml OP-1 (6C) or no morphogen (6D). Figs. 6C and 6D are at 10x and 20x magnification compared to FIGs 6A and 5B. As can readily be seen, in the presence of OP-1, EC cells grow as flattened cells, becoming anchorage dependent. In addition, growth rate is reduced approximately 10 fold. Finally, the cells are induced to differentiate.

35

MAINTENANCE OF PHENOTYPE

The morphogens of this invention also may be used to maintain a cell's differentiated phenotype. This morphogenic capability is particularly useful for  
5 inducing the continued expression of phenotype in senescent or quiescent cells.

The phenotypic maintenance capability of morphogens is readily assessed. A number of  
10 differentiated cells become senescent or quiescent after multiple passages under standard tissue culture conditions in vitro. However, if these cells are cultivated in vitro in association with a morphogen of this invention, the cells are induced to maintain  
15 expression of their phenotype through multiple passages. For example, the alkaline phosphatase activity of cultured osteoblasts, like cultured osteosarcoma cells and calvaria cells, is significantly reduced after multiple passages in vitro.  
20 However, if the cells are cultivated in the presence of a morphogen (e.g., OP-1), alkaline phosphatase activity is maintained over extended periods of time. Similarly, phenotypic expression of myocytes also is maintained in the presence of the morphogen. This  
25 experiment may be performed with other morphogens and different cells to assess the phenotypic maintenance capability of different morphogens on cells of differing origins.

30 Phenotypic maintenance capability also may be assessed in vivo, using a rat model for osteoporosis, disclosed in co-pending USSN 752,857, filed August 30, 1991,, incorporated herein by reference. As disclosed therein, Long Evans rats are ovariectomized to produce  
35 an osteoporotic condition resulting from decreased

estrogen production. Eight days after ovariectomy, rats are systemically provided with phosphate buffered saline (PBS) or OP-1 (21  $\mu$ g or 20  $\mu$ g) for 22 days. The rats then are sacrificed and serum alkaline phosphatase levels, serum calcium levels, and serum osteocalcin levels are determined, using standard methodologies. Three-fold higher levels of osteocalcin levels are found in rats provided with 1 or 20  $\mu$ g of OP-1. Increased alkaline phosphatase levels also were seen. Histomorphometric analysis on the tibial diaphysical bone shows OP-1 can reduce bone mass lost due to the drop in estrogen levels.

#### CELL STIMULATION

The ability of the morphogens of this invention to stimulate the proliferation of progenitor cells also can be assayed readily in vitro. Useful naive stem cells include pluripotential stem cells, which may be isolated from bone marrow or umbilical cord blood using conventional methodologies, (see, for example, Faradji et al., (1988) Vox Sang. 55 (3):133-138 or Broxmeyer et al., (1989) PNAS 86 (10):3828-3832), as well as naive stem cells obtained from blood. Alternatively, embryonic cells (e.g., from a cultured mesodermal cell line) may be useful.

Another method for obtaining progenitor cells and for determining the ability of morphogens to stimulate cell proliferation is to capture progenitor cells from an in vivo source. For example, a biocompatible matrix material able to allow the influx of migratory progenitor cells may be implanted at an in vivo site long enough to allow the influx of migratory progenitor cells. For example, a bone-derived,

guanidine-extracted matrix, formulated as disclosed for example in Sampath et al. ((1983) PNAS 80:6591-6595), or U.S. Patent No. 4,975,526, may be implanted into a rat at a subcutaneous site, essentially following the method of Sampath et al. (ibid). After three days the implant is removed, and the progenitor cells associated with the matrix dispersed and cultured.

Progenitor cells, however obtained, then are incubated in vitro with a suspected morphogen under standard cell culture conditions well known to those having ordinary skill in the art. In the absence of external stimuli, the progenitor cells do not, or minimally proliferate on their own in culture. However, if the cells are cultured in the presence of a morphogen, such as OP-1, they are stimulated to proliferate. Cell growth can be determined visually or spectrophotometrically using standard methods well known in the art.

20

#### PROLIFERATION OF PROGENITOR CELL POPULATIONS

Progenitor cells may be stimulated to proliferate in vivo or ex vivo. The cells may be stimulated in vivo by injecting or otherwise providing a sterile preparation containing the morphogen into the individual. For example, the hemopoietic pluripotential stem cell population of an individual may be stimulated to proliferate by injecting or otherwise providing an appropriate concentration of the morphogen to the individual's bone marrow.

Progenitor cells may be stimulated ex vivo by contacting progenitor cells of the population to be enhanced with a morphogen under sterile conditions at a

concentration and for a time sufficient to stimulate proliferation of the cells. In general, a period of from about 10 minutes to about 24 hours should be sufficient. The stimulated cells then are provided to the individual as, for example, by injecting the cells to an appropriate in vivo locus. Suitable biocompatible progenitor cells may be obtained by any of the methods known in the art or described herein.

10           REGENERATION OF DAMAGED OR DISEASED TISSUE

          The morphogens of this invention may be used to repair diseased or damaged mammalian tissue. The tissue to be repaired is preferably assessed, and  
15 excess necrotic or interfering scar tissue removed as needed, by surgical, chemical, ablating or other methods known in the medical arts.

          The morphogen then may be provided directly to  
20 the tissue locus as part of a sterile, biocompatible composition, either by surgical implantation or injection. Alternatively, a sterile, biocompatible composition containing morphogen-stimulated progenitor cells may be provided to the tissue locus. The  
25 existing tissue at the locus, whether diseased or damaged, provides the appropriate matrix to allow the proliferation and tissue-specific differentiation of progenitor cells. In addition, a damaged or diseased tissue locus, particularly one that has been further  
30 assaulted by surgical means, provides a morphogenically permissive environment. For some tissues, it is envisioned that systemic provision of the morphogen will be sufficient.

In some circumstances, particularly where tissue damage is extensive, the tissue may not be capable of providing a sufficient matrix for cell influx and proliferation. In these instances, it may  
5 be necessary to provide the morphogen or morphogen-stimulated progenitor cells to the tissue locus in association with a suitable, biocompatible formulated matrix, prepared by any of the means described below. The matrix preferably is tissue-specific, in vivo  
10 biodegradable, and comprises particles having dimensions within the range of 70-850 $\mu$ m, most preferably 150-420 $\mu$ m.

The morphogens of this invention also may be used  
15 to prevent or substantially inhibit scar tissue formation following an injury. If a morphogen is provided to a newly injured tissue locus, it can induce tissue morphogenesis at the locus, preventing the aggregation of migrating fibroblasts into non-  
20 differentiated connective tissue. The morphogen preferably is provided as a sterile pharmaceutical preparation injected into the tissue locus within five hours of the injury. Several non-limiting examples follow, illustrating the morphogens regenerate  
25 capabilities in different tissues. The proteins of this invention previously have been shown to be capable of inducing cartilage and endochondral bone formation (See, for example U.S. Patent No. 5,011,691).

30 As an example, protein-induced morphogenesis of substantially injured liver tissue following a partial hepatectomy is disclosed. Variations on this general protocol may be used to test morphogen activity in other different tissues. The general method involves  
35 excising an essentially nonregenerating portion of a



tissue and providing the morphogen, preferably as a soluble pharmaceutical preparation to the excised tissue locus, closing the wound and examining the site at a future date. Like bone, liver has a potential to  
5 regenerate upon injury during post-fetal life.

Morphogen, (e.g., purified recombinant human OP-1, mature form), was solubilized (1 mg/ml) in 50% ethanol (or compatible solvent) containing 0.1% trifluoroacetic  
10 acid (or compatible acid). The injectable OP-1 solution was prepared by diluting one volume of OP-1/solvent-acid stock solution with 9 volumes of 0.2% rat serum albumin in sterile PBS (phosphate-buffered saline).

15

Growing rats or aged rats were anesthetized by using ketamine. Two of the liver lobes (left and right) were cut out (approximately 1/3 of the lobe) and the OP-1 was injected locally at multiple sites along  
20 the cut ends. The amount of OP-1 injected was 100  $\mu$ g in 100 of PBS/RSA (phosphate buffered saline/rat serum albumin) injection buffer. Placebo samples are injection buffer without OP-1. Five rats in each group were used. The wound was closed and the rats were  
25 allowed to eat normal food and drink tap water.

After 12 days, the rats were sacrificed and liver regeneration was observed visually. The photomicrograph in Fig. 7 illustrates dramatically the  
30 regenerative effects of OP-1 on liver regeneration. The OP-1-injected group showed complete liver tissue regeneration and no sign remained of any cut in the liver (animal 2). By contrast, in the control group into which only PBS was injected only minimal  
35 regeneration was evidenced (animal 1). In addition, the incision remains in this sample.

As another example, the ability of the morphogens of this invention to induce dentinogenesis also was assessed. To date, the unpredictable response of dental pulp tissue to injury is a basic clinical problem in dentistry. Cynomolgus monkeys were chosen as primate models as monkeys are presumed to be more indicative of human dental biology than models based on lower non-primate mammals.

10

Using standard dental surgical procedures, small areas (e.g., 2mm) of dental pulps were surgically exposed by removing the enamel and dentin immediately above the pulp (by drilling) of sample teeth, performing a partial amputation of the coronal pulp tissue, inducing hemostasis, application of the pulp treatment, and sealing and filling the cavity by standard procedures.

20 Pulp treatments used were: OP-1 dispersed in a carrier matrix; carrier matrix alone and no treatment. Twelve teeth per animal (four for each treatment) were prepared, and two animals were used. At four weeks, teeth were extracted and processed histologically for analysis of dentin formation, and/or ground to analyze dentin mineralization. FIG.8 illustrates dramatically the effect of morphogen on osteodentin reparation. FIG. 8A is a photomicrograph of the control treatment (PBS) and shows little or no reparation. FIG. 8B is a photomicrograph of treatment with carrier alone, showing minimal reparation. By contrast, treatment with morphogen (FIG. 8C) shows significant reparation. The results of FIG. 8 indicate that OP-1-CM (OP-1 plus

25

30

carrier matrix) reliably induced formation of reparative or osteodentin bridges on surgically exposed healthy dental pulps. By contrast, pulps treated with carrier matrix alone, or not treated failed to form  
5 reparative dentin.

As another example, the morphogen-induced regenerative effects on central nervous system (CNS) repair may be assessed using a rat brain stab model.  
10 Briefly, male Long Evans rats are anesthetized and the head area prepared for surgery. The calvariae is exposed using standard surgical procedures and a hole drilled toward the center of each lobe using a 0.035K wire, just piercing the calvariae. 25 $\mu$ l solutions  
15 containing either morphogen (OP-1, 25 $\mu$ g) or PBS then is provided to each of the holes by Hamilton syringe. Solutions are delivered to a depth approximately 3 mm below the surface, into the underlying cortex, corpus callosum and hippocampus. The skin then is sutured and  
20 the animal allowed to recover.

Three days post surgery, rats are sacrificed by decapitation and their brains processed for sectioning. Scar tissue formation is evaluated by immunofluorescence staining for glial fibrillary acidic protein, a marker  
25 protein for glial scarring, to qualitatively determine the degree of scar formation. Sections also are probed with anti-OP-1 antibodies to determine the presence of OP-1. Reduced levels of glial fibrillary acidic  
30 protein are anticipated in the tissue sections of animals treated with morphogen, evidencing the ability of morphogen to inhibit glial scar formation, thereby stimulating nerve regeneration.

Antibodies to morphogens of this invention have been identified in healthy human sera. In addition, implanting devices comprising morphogens (e.g., OP-1) have been discovered to induce an increase in anti-morphogen antibodies (e.g., anti-OP-1 antibodies). It is anticipated that these antibodies comprise part of the body's regulation of morphogen activity in vivo. The presence of the antibodies, and fluctuations in their levels, which are readily monitored, can provide a useful method for monitoring tissue stasis and tissue viability (e.g., identification of a pathological state). For example, standard radioimmunoassays or ELISA may be used to detect and quantify endogenous anti-morphogen antibodies in sera. Antibodies or other binding proteins capable of detecting anti-morphogen antibodies may be obtained using standard methodologies.

20

#### MATRIX PREPARATION

The morphogens of this invention may be implanted surgically, dispersed in a biocompatible, preferably in vivo biodegradable matrix appropriately modified to provide a structure in which the morphogen may be dispersed and which allows the influx, differentiation and proliferation of migrating progenitor cells. The matrix also should provide signals capable of directing the tissue specificity of the differentiating cells, as well as a morphogenically permissive environment, being essentially free of growth inhibiting signals.

In the absence of these features the matrix does not appear to be suitable as part of a morphogenic composition. Recent studies on osteogenic devices

(morphogens dispersed within a formulated matrix) using matrices formulated from polylactic acid and/or polyglycolic acid biopolymers, ceramics ( $\alpha$ -tri-calcium-phosphate), or hydroxyapatite show that these materials, by themselves, are unable to provide the appropriate environment for inducing de novo endochondral bone formation in rats by themselves. In addition, matrices formulated from commercially available highly purified, reconstituted collagens or naturally-derived non-bone, species-specific collagen (e.g., from rat tail tendon) also are unsuccessful in inducing bone when implanted in association with an osteogenic protein. These matrices apparently lack specific structurally-related features which aid in directing the tissue specificity of the morphogen-stimulated, differentiating progenitor cells.

The formulated matrix may be shaped as desired in anticipation of surgery or may be shaped by the physician or technician during surgery. Thus, the material may be used in topical, subcutaneous, intraperitoneal, or intramuscular implants to repair tissue or to induce its growth de novo. The matrix preferably is biodegradable in vivo, being slowly absorbed by the body and replaced by new tissue growth, in the shape or very nearly in the shape of the implant.

Details of how to make and how to use the matrices useful in this invention are disclosed below.

#### TISSUE-DERIVED MATRICES

Suitable biocompatible, in vivo biodegradable acellular matrices may be prepared from naturally-

occurring tissue. The tissue is treated with suitable agents to substantially extract the cellular, nonstructural components of the tissue. The agents also should be capable of extracting any growth  
5 inhibiting components associated with the tissue. The resulting material is a porous, acellular matrix, substantially depleted in nonstructurally-associated components.

10           The matrix also may be further treated with agents that modify the matrix, increasing the number of pores and micropits on its surfaces. Those skilled in the art will know how to determine which agents are best suited to the extraction of nonstructural  
15 components for different tissues. For example, soft tissues such as liver and lung may be thin-sectioned and exposed to a nonpolar solvent such as, for example, 100% ethanol, to destroy the cellular structure of the tissue and extract nonstructural components. The  
20 material then is dried and pulverized to yield nonadherent porous particles. Structural tissues such as cartilage and dentin where collagen is the primary component may be demineralized and extracted with guanidine, essentially following the method of Sampath  
25 et al. (1983) PNAS 80:6591-6595. For example, pulverized and demineralized dentin is extracted with five volumes of 4M guanidine-HCl, 50mM Tris-HCl, pH 7.0 for 16 hours at 4°C. The suspension then is filtered. The insoluble material that remains is collected and  
30 used to fabricate the matrix. The material is mostly collagenous in manner. It is devoid of morphogenic activity. The matrix particles may further be treated with a collagen fibril-modifying agent that extracts potentially unwanted components from the matrix, and  
35 alters the surface structure of the matrix material.

Useful agents include acids, organic solvents or heated aqueous media. A detailed description of these matrix treatments are disclosed in U.S. Patent No. 4,975,526 and PCT publication US90/00912, published September 7, 1990 (WO90/10018).

The currently most preferred agent is a heated aqueous fibril-modifying medium such as water, to increase the matrix particle surface area and porosity.

10 The currently most preferred aqueous medium is an acidic aqueous medium having a pH of less than about 4.5, e.g., within the range of about pH 2 - pH 4 which may help to "swell" the collagen before heating. 0.1% acetic acid, which has a pH of about 3, currently is

15 most preferred. 0.1 M acetic acid also may be used.

Various amounts of delipidated, demineralized guanidine-extracted bone collagen are heated in the aqueous medium (1g matrix/30ml aqueous medium) under

20 constant stirring in a water jacketed glass flask, and maintained at a given temperature for a predetermined period of time. Preferred treatment times are about one hour, although exposure times of between about 0.5 to two hours appear acceptable. The temperature

25 employed is held constant at a temperature within the range of about 37°C to 65°C. The currently preferred heat treatment temperature is within the range of about 45°C to 60°C.

30 After the heat treatment, the matrix is filtered, washed, lyophilized and used for implant. Where an acidic aqueous medium is used, the matrix also is preferably neutralized prior to washing and lyophilization. A currently preferred neutralization

35 buffer is a 200mM sodium phosphate buffer, pH 7.0. To

neutralize the matrix, the matrix preferably first is allowed to cool following thermal treatment, the acidic aqueous medium (e.g., 0.1% acetic acid) then is removed and replaced with the neutralization buffer and the  
5 matrix agitated for about 30 minutes. The neutralization buffer then may be removed and the matrix washed and lyophilized.

Other useful fibril-modifying treatments include  
10 acid treatments (e.g., trifluoroacetic acid and hydrogen fluoride) and solvent treatments such as dichloromethane, acetonitrile, isopropanol and chloroform, as well as particular acid/solvent combinations.

15

After contact with the fibril-modifying agent, the treated matrix may be washed to remove any extracted components, following a form of the procedure set forth below:

20

1. Suspend matrix preparation in TBS (Tris-buffered saline) 1g/200 ml and stir at 4°C for 2 hrs; or in 6 M urea, 50 mM Tris-HCl, 500 mM NaCl, pH 7.0 (UTBS) or water and stir at room temperature (RT) for  
25 30 minutes (sufficient time to neutralize the pH);

2. Centrifuge and repeat wash step; and

3. Centrifuge; discard supernatant; water  
30 wash residue; and then lyophilize.

#### SYNTHETIC TISSUE-SPECIFIC MATRICES

35

In addition to the naturally-derived tissue-



specific matrices described above, useful tissue-specific matrices may be formulated synthetically if appropriately modified. These porous biocompatible, in vivo biodegradable synthetic matrices are disclosed in

5 PCT publication US91/03603, published December 12, 1991 (WO91/18558), the disclosure of which is hereby incorporated by reference. Briefly, the matrix comprises a porous crosslinked structural polymer of biocompatible, biodegradable collagen and appropriate,

10 tissue-specific glycosaminoglycans as tissue-specific cell attachment factors. Collagen derived from a number of sources may be suitable for use in these synthetic matrices, including insoluble collagen, acid-soluble collagen, collagen soluble in neutral or basic

15 aqueous solutions, as well as those collagens which are commercially available.

Glycosaminoglycans (GAGs) or mucopolysaccharides are hexosamine-containing

20 polysaccharides of animal origin that have a tissue specific distribution, and therefore may be used to help determine the tissue specificity of the morphogen-stimulated differentiating cells. Reaction with the GAGs also provides collagen with another valuable

25 property, i.e., inability to provoke an immune reaction (foreign body reaction) from an animal host.

Chemically, GAGs are made up of residues of hexoseamines glycosidically bound and alternating in a

30 more-or-less regular manner with either hexouronic acid or hexose moieties (see, e.g., Dodgson et al. in Carbohydrate Metabolism and its Disorders (Dickens et al., eds.) Vol. 1, Academic Press (1968)). Useful GAGs include hyaluronic acid, heparin, heparin sulfate,

35 chondroitin 6-sulfate, chondroitin 4-sulfate, dermatan

sulfate, and keratin sulfate. Other GAGs are suitable for forming the matrix described herein, and those skilled in the art will either know or be able to ascertain other suitable GAGs using no more than  
5 routine experimentation. For a more detailed description of mucopolysaccharides, see Aspinall, Polysaccharides, Pergamon Press, Oxford (1970). For example, as disclosed in U.S. Application Serial No. 529,852, chondroitin-6-sulfate can be used where  
10 endochondral bone formation is desired. Heparin sulfate, on the other hand, may be used to formulate synthetic matrices for use in lung tissue repair.

Collagen can be reacted with a GAG in aqueous  
15 acidic solutions, preferably in diluted acetic acid solutions. By adding the GAG dropwise into the aqueous collagen dispersion, coprecipitates of tangled collagen fibrils coated with GAG results. This tangled mass of fibers then can be homogenized to form a homogeneous  
20 dispersion of fine fibers and then filtered and dried.

Insolubility of the collagen-GAG products can be raised to the desired degree by covalently cross-linking these materials, which also serves to raise the  
25 resistance to resorption of these materials. In general, any covalent cross-linking method suitable for cross-linking collagen also is suitable for cross-linking these composite materials, although crosslinking by a dehydrothermal process is preferred.

30 When dry, the crosslinked particles are essentially spherical, with diameters of about 500  $\mu\text{m}$ . Scanning electron microscopy shows pores of about 20  $\mu\text{m}$  on the surface and 40  $\mu\text{m}$  on the interior. The  
35 interior is made up of both fibrous and sheet-like

structures, providing surfaces for cell attachment.  
The voids interconnect, providing access to the cells  
throughout the interior of the particle. The material  
appears to be roughly 99.5% void volume, making the  
5 material very efficient in terms of the potential cell  
mass that can be grown per gram of microcarrier.

The morphogens described herein can be  
combined and dispersed in an appropriately modified  
10 tissue-specific matrix using any of the methods  
described below:

1. Ethanol Precipitation

15 Matrix is added to the morphogen dissolved in  
guanidine-HCl. Samples are vortexed and incubated at a  
low temperature. Samples are then further vortexed.  
Cold absolute ethanol is added to the mixture which is  
then stirred and incubated. After centrifugation  
20 (microfuge, high speed) the supernatant is discarded.  
The matrix is washed with cold concentrated ethanol in  
water and then lyophilized.

2. Acetonitrile Trifluoroacetic  
25 Acid Lyophilization

In this procedure, morphogen in an  
acetonitrile trifluoroacetic acid (ACN/TFA solution is  
added to the carrier material. Samples are vigorously  
30 vortexed many times and then lyophilized.

3. Buffered Saline Lyophilization

Morphogen preparations in physiological saline  
35 may also be vortexed with the matrix and lyophilized to

produce morphogenically active material.

#### BIOASSAY

5

The following sets forth various procedures for evaluating the in vivo morphogenic utility of the morphogens and morphogenic compositions of this invention. The proteins and compositions may be  
10 injected or surgically implanted in a mammal, following any of a number of procedures well known in the art. For example, surgical implant bioassays may be performed essentially following the procedure of Sampath et al. (1983) PNAS 80:6591-6595.

15

#### Histological Evaluation

Histological sectioning and staining is preferred to determine the extent of morphogenesis in vivo, particularly in tissue repair procedures.  
20 Excised implants are fixed in Bouins Solution, embedded in paraffin, and cut into 6-8  $\mu$ m sections. Staining with toluidine blue or hemotoxylin/eosin demonstrates clearly the ultimate development of the new tissue.  
25 Twelve day implants are usually sufficient to determine whether the implants contain newly induced tissue.

Successful implants exhibit a controlled progression through the stages of induced tissue  
30 development allowing one to identify and follow the tissue-specific events that occur. For example, in endochondral bone formation the stages include:  
(1) leukocytes on day one; (2) mesenchymal cell migration and proliferation on days two and three;  
35 (3) chondrocyte appearance on days five and six;

- (4) cartilage matrix formation on day seven;  
(5) cartilage calcification on day eight; (6) vascular invasion, appearance of osteoblasts, and formation of new bone on days nine and ten; (7) appearance of  
5 osteoblastic and bone remodeling and dissolution of the implanted matrix on days twelve to eighteen; and  
(8) hematopoietic bone marrow differentiation in the ossicle on day twenty-one.

10 Biological Markers

- In addition to histological evaluation, biological markers may be used as a marker for tissue morphogenesis. Useful markers include tissue-specific  
15 enzymes whose activities may be assayed (e.g., spectrophotometrically) after homogenization of the implant. These assays may be useful for quantitation and for obtaining an estimate of tissue formation quickly after the implants are removed from the animal.  
20 For example, alkaline phosphatase activity may be used as a marker for osteogenesis.

- Incorporation of systemically provided morphogens may be followed using tagged morphogens  
25 (e.g., radioactively labelled) and determining their localization in new tissue, and/or by monitoring their disappearance from the circulatory system using a standard pulse-chase labeling protocol. The morphogen also may be provided with a tissue-specific molecular  
30 tag, whose uptake may be monitored and correlated with the concentration of morphogen provided. As an example, ovary removal in female rats results in reduced bone alkaline phosphatase activity, rendering the rats predisposed to osteoporosis. If the female  
35 rats now are provided with a morphogen, e.g., OP-1, a

reduction in the systemic concentration of calcium ( $CA^{2+}$ ) is seen, which correlates with the presence of the provided morphogen and can be shown to correspond to increased alkaline phosphatase activity.

5

The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The present embodiments are therefore to be considered in all respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced therein.

15

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: COHEN, CHARLES M.  
KUBERASAMPATH, THANGAVEL  
PANG, ROY H.L.  
OPPERMANN, HERMANN  
RUEGER, DAVID C.

(ii) TITLE OF INVENTION: PROTEIN-INDUCED MORPHOGENESIS

(iii) NUMBER OF SEQUENCES: 23

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
(B) STREET: 53 STATE STREET  
(C) CITY: BOSTON  
(D) STATE: MASSACHUSETTS  
(E) COUNTRY: U.S.A.  
(F) ZIP: 02109

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 667,274  
(B) FILING DATE: 11-MAR-1991

## (vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 752,764  
(B) FILING DATE: 30-AUG-1991

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 97 amino acids  
(B) TYPE: amino acids  
(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME: Generic Sequence 1

(D) OTHER INFORMATION: Each Xaa indicates one of the 20 naturally-occurring L-isomer,  $\alpha$ -amino acids or a derivative thereof.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

Xaa Xaa Xaa Xaa Xaa Xaa
  1                               5
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 10                               15
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa
 20                               25
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 30                               35
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 40                               45                               50
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
                               55                               60
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                               65                               70
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                               75                               80
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
                               85                               90
Xaa Cys Xaa
                               95
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 97 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear



(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME: Generic Sequence 2

(D) OTHER INFORMATION: Each Xaa indicates one of the 20 naturally-occurring L-isomer,  $\alpha$ -amino acids or a derivative thereof.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

                                Xaa Xaa Xaa Xaa Xaa Xaa
                                1           5
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                                10          15
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa
                                20          25
Cys Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa
                                30          35
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                                40          45          50
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
                                55          60
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                                65          70
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                                75          80
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
                                85          90
Xaa Cys Xaa
                                95

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 97 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
- (ix) FEATURE:
- (A) NAME: Generic Sequence 3
- (D) OTHER INFORMATION: wherein each Xaa is independently selected from a group of one or more specified amino acids as defined in the specification.
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Leu Tyr Val Xaa Phe
 1             5
Xaa Xaa Xaa Gly Trp Xaa Xaa Trp Xaa
             10
Xaa Ala Pro Gly Xaa Xaa Xaa Ala
 15             20
Xaa Tyr Cys Xaa Gly Xaa Cys Xaa
             25             30
Xaa Pro Xaa Xaa Xaa Xaa Xaa
             35
Xaa Xaa Xaa Asn His Ala Xaa Xaa
             40             45
Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa
             50
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys
             55             60
Cys Xaa Pro Xaa Xaa Xaa Xaa Xaa
             65
Xaa Xaa Xaa Leu Xaa Xaa Xaa
             70             75
Xaa Xaa Xaa Xaa Val Xaa Leu Xaa
             80
Xaa Xaa Xaa Xaa Met Xaa Val Xaa
             85             90
Xaa Cys Gly Cys Xaa
             95
```

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (ix) FEATURE:

(A) NAME: Generic Sequence 4

(D) OTHER INFORMATION: wherein each Xaa is independently selected from a group of one or more specified amino acids as defined in the specification.

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Cys	Xaa	Xaa	Xaa	Xaa	Leu	Tyr	Val	Xaa	Phe
1				5					10
Xaa	Xaa	Xaa	Gly	Trp	Xaa	Xaa	Trp	Xaa	
				15					
Xaa	Ala	Pro	Xaa	Gly	Xaa	Xaa	Ala		
20					25				
Xaa	Tyr	Cys	Xaa	Gly	Xaa	Cys	Xaa		
		30					35		
Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Xaa			
				40					
Asn	Xaa	Xaa	Asn	His	Ala	Xaa	Xaa		
		45					50		
Xaa	Xaa	Leu	Xaa	Xaa	Xaa	Xaa	Xaa		
				55					
Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Xaa	Cys		
		60					65		
Cys	Xaa	Pro	Xaa	Xaa	Xaa	Xaa	Xaa		
				70					

```

Xaa Xaa Xaa Leu Xaa Xaa Xaa
 75                               80
Xaa Xaa Xaa Xaa Val Xaa Leu Xaa
                        85
Xaa Xaa Xaa Xaa Met Xaa Val Xaa
 90                               95
Xaa Cys Gly Cys Xaa
                        100

```

- (2) INFORMATION FOR SEQ ID NO:5:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 139 amino acids
    - (B) TYPE: amino acids
    - (C) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (ix) FEATURE:
    - (A) NAME: hOP-1 (mature form)
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```

Ser Thr Gly Ser Lys Gln Arg Ser Gln
 1                               5
Asn Arg Ser Lys Thr Pro Lys Asn Gln
10                               15
Glu Ala Leu Arg Met Ala Asn Val Ala
20                               25
Glu Asn Ser Ser Ser Asp Gln Arg Gln
30                               35
Ala Cys Lys Lys His Glu Leu Tyr Val
40                               45
Ser Phe Arg Asp Leu Gly Trp Gln Asp
50
Trp Ile Ile Ala Pro Glu Gly Tyr Ala
55                               60
Ala Tyr Tyr Cys Glu Gly Glu Cys Ala
65                               70

```

Phe	Pro	Leu	Asn	Ser	Tyr	Met	Asn	Ala
		75					80	
Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu
			85					90
Val	His	Phe	Ile	Asn	Pro	Glu	Thr	Val
				95				
Pro	Lys	Pro	Cys	Cys	Ala	Pro	Thr	Gln
100					105			
Leu	Asn	Ala	Ile	Ser	Val	Leu	Tyr	Phe
	110					115		
Asp	Asp	Ser	Ser	Asn	Val	Ile	Leu	Lys
		120					125	
Lys	Tyr	Arg	Asn	Met	Val	Val	Arg	Ala
			130					135
Cys	Gly	Cys	His					

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 139 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME: mOP-1 (mature form)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ser	Thr	Gly	Gly	Lys	Gln	Arg	Ser	Gln
1				5				
Asn	Arg	Ser	Lys	Thr	Pro	Lys	Asn	Gln
10					15			
Glu	Ala	Leu	Arg	Met	Ala	Ser	Val	Ala
	20					25		

Glu	Asn	Ser	Ser	Ser	Asp	Gln	Arg	Gln
		30					35	
Ala	Cys	Lys	Lys	His	Glu	Leu	Tyr	Val
			40					45
Ser	Phe	Arg	Asp	Leu	Gly	Trp	Gln	Asp
				50				
Trp	Ile	Ile	Ala	Pro	Glu	Gly	Tyr	Ala
55					60			
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ala
	65					70		
Phe	Pro	Leu	Asn	Ser	Tyr	Met	Asn	Ala
		75					80	
Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu
			85					90
Val	His	Phe	Ile	Asn	Pro	Asp	Thr	Val
				95				
Pro	Lys	Pro	Cys	Cys	Ala	Pro	Thr	Gln
100					105			
Leu	Asn	Ala	Ile	Ser	Val	Leu	Tyr	Phe
	110					115		
Asp	Asp	Ser	Ser	Asn	Val	Ile	Leu	Lys
		120					125	
Lys	Tyr	Arg	Asn	Met	Val	Val	Arg	Ala
			130					135
Cys	Gly	Cys	His					

- (2) INFORMATION FOR SEQ ID NO:7:
- (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 139 amino acids
    - (B) TYPE: amino acids
    - (C) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein
  - (ix) FEATURE:
    - (A) NAME: hOP-2 (mature form)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ala	Val	Arg	Pro	Leu	Arg	Arg	Arg	Gln
1				5				
Pro	Lys	Lys	Ser	Asn	Glu	Leu	Pro	Gln
10					15			
Ala	Asn	Arg	Leu	Pro	Gly	Ile	Phe	Asp
20						25		
Asp	Val	His	Gly	Ser	His	Gly	Arg	Gln
		30					35	
Val	Cys	Arg	Arg	His	Glu	Leu	Tyr	Val
			40					45
Ser	Phe	Gln	Asp	Leu	Gly	Trp	Leu	Asp
				50				
Trp	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser
55					60			
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ser
65						70		
Phe	Pro	Leu	Asp	Ser	Cys	Met	Asn	Ala
		75					80	
Thr	Asn	His	Ala	Ile	Leu	Gln	Ser	Leu
			85					90
Val	His	Leu	Met	Lys	Pro	Asn	Ala	Val
				95				
Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys
100					105			
Leu	Ser	Ala	Thr	Ser	Val	Leu	Tyr	Tyr
110						115		
Asp	Ser	Ser	Asn	Asn	Val	Ile	Leu	Arg
		120					125	
Lys	His	Arg	Asn	Met	Val	Val	Lys	Ala
			130					135
Cys	Gly	Cys	His					

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 139 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (ix) FEATURE:

(A) NAME: mOP-2 (mature form)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ala	Ala	Arg	Pro	Leu	Lys	Arg	Arg	Gln
1				5				
Pro	Lys	Lys	Thr	Asn	Glu	Leu	Pro	His
10					15			
Pro	Asn	Lys	Leu	Pro	Gly	Ile	Phe	Asp
20						25		
Asp	Gly	His	Gly	Ser	Arg	Gly	Arg	Glu
		30					35	
Val	Cys	Arg	Arg	His	Glu	Leu	Tyr	Val
			40					45
Arg	Phe	Arg	Asp	Leu	Gly	Trp	Leu	Asp
				50				
Trp	Val	Ile	Ala	Pro	Gln	Gly	Tyr	Ser
55					60			
Ala	Tyr	Tyr	Cys	Glu	Gly	Glu	Cys	Ala
65						70		
Phe	Pro	Leu	Asp	Ser	Cys	Met	Asn	Ala
		75					80	
Thr	Asn	His	Ala	Ile	Leu	Gln	Ser	Leu
			85					90
Val	His	Leu	Met	Lys	Pro	Asp	Val	Val
				95				
Pro	Lys	Ala	Cys	Cys	Ala	Pro	Thr	Lys
100					105			



Leu	Ser	Ala	Thr	Ser	Val	Leu	Tyr	Tyr
	110					115		
Asp	Ser	Ser	Asn	Asn	Val	Ile	Leu	Arg
		120					125	
Lys	His	Arg	Asn	Met	Val	Val	Lys	Ala
			130					135
Cys	Gly	Cys	His					

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 96 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME: CBMP-2A(fx)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Cys	Lys	Arg	His	Pro	Leu	Tyr	Val	Asp	Phe	Ser
1				5					10	
Asp	Val	Gly	Trp	Asn	Asp	Trp	Ile	Val	Ala	Pro
		15						20		
Pro	Gly	Tyr	His	Ala	Phe	Tyr	Cys	His	Gly	Glu
	25					30				
Cys	Pro	Phe	Pro	Leu	Ala	Asp	His	Leu	Asn	Ser
	35					40				
Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	Asn
	45				50				55	
Ser	Val	Asn	Ser	Lys	Ile	Pro	Lys	Ala	Cys	Cys
		60						65		
Val	Pro	Thr	Glu	Leu	Ser	Ala	Ile	Ser	Met	Leu
	70						75			
Tyr	Leu	Asp	Glu	Asn	Glu	Lys	Val	Val	Leu	Lys
	80						85			

```

Asn Tyr Gln Asp Met Val Val Glu Gly Cys Gly
      90                      95
Cys Arg
100

```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 101 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(ix) **FEATURE:**

(A) NAME: CBMP-2B(fx)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

						Cys	Arg	Arg	His	Ser	
						1					5
Leu	Tyr	Val	Asp	Phe	Ser	Asp	Val	Gly	Trp	Asn	
				10						15	
Asp	Trp	Ile	Val	Ala	Pro	Pro	Gly	Tyr	Gln	Ala	
			20						25		
Phe	Tyr	Cys	His	Gly	Asp	Cys	Pro	Phe	Pro	Leu	
			30						35		
Ala	Asp	His	Leu	Asn	Ser	Thr	Asn	His	Ala	Ile	
			40						45		
Val	Gln	Thr	Leu	Val	Asn	Ser	Val	Asn	Ser	Ser	
			50			55				60	
Ile	Pro	Lys	Ala	Cys	Cys	Val	Pro	Thr	Glu	Leu	
				65						70	
Ser	Ala	Ile	Ser	Met	Leu	Tyr	Leu	Asp	Glu	Tyr	
			75							80	
Asp	Lys	Val	Val	Leu	Lys	Asn	Tyr	Gln	Glu	Met	
			85							90	

Val.	Val	Glu	Gly	Cys	Gly	Cys	Arg
	95					100	

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) **FEATURE:**

(A) NAME: DPP(fx)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser
 1                     5                    10
Asp Val Gly Trp Asp Asp Trp Ile Val Ala Pro
          15                      20
Leu Gly Tyr Asp Ala Tyr Tyr Cys His Gly Lys
        25                      30
Cys Pro Phe Pro Leu Ala Asp His Phe Asn Ser
      35                      40
Thr Asn His Ala Val Val Gln Thr Leu Val Asn
    45                      50                    55
Asn Asn Asn Pro Gly Lys Val Pro Lys Ala Cys
            60                      65
Cys Val Pro Thr Gln Leu Asp Ser Val Ala Met
              70                      75
Leu Tyr Leu Asn Asp Gln Ser Thr Val Val Leu
           80                      85
Lys Asn Tyr Gln Glu Met Thr Val Val Gly Cys
       90                      95
Gly Cys Arg
100
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(A) NAME: Vgl(fx)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Cys Lys Lys Arg His Leu Tyr Val Glu Phe Lys  
1 5 10  
Asp Val Gly Trp Gln Asn Trp Val Ile Ala Pro  
15 20  
Gln Gly Tyr Met Ala Asn Tyr Cys Tyr Gly Glu  
25 30  
Cys Pro Tyr Pro Leu Thr Glu Ile Leu Asn Gly  
35 40  
Ser Asn His Ala Ile Leu Gln Thr Leu Val His  
45 50 55  
Ser Ile Glu Pro Glu Asp Ile Pro Leu Pro Cys  
60 65  
Cys Val Pro Thr Lys Met Ser Pro Ile Ser Met  
70 75  
Leu Phe Tyr Asp Asn Asn Asp Asn Val Val Leu  
80 85  
Arg His Tyr Glu Asn Met Ala Val Asp Glu Cys  
90 95  
Gly Cys Arg  
100

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids

(B) TYPE: amino acids

(C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (ix) FEATURE:

(A) NAME: Vgr-1(fx)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Cys	Lys	Lys	His	Glu	Leu	Tyr	Val	Ser	Phe	Gln										
1					5					10										
Asp	Val	Gly	Trp	Gln	Asp	Trp	Ile	Ile	Ala	Pro										
				15					20											
Xaa	Gly	Tyr	Ala	Ala	Asn	Tyr	Cys	Asp	Gly	Glu										
				25				30												
Cys	Ser	Phe	Pro	Leu	Asn	Ala	His	Met	Asn	Ala										
				35				40												
Thr	Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	His										
				45				50		55										
Val	Met	Asn	Pro	Glu	Tyr	Val	Pro	Lys	Pro	Cys										
				60					65											
Cys	Ala	Pro	Thr	Lys	Val	Asn	Ala	Ile	Ser	Val										
				70				75												
Leu	Tyr	Phe	Asp	Asp	Asn	Ser	Asn	Val	Ile	Leu										
				80				85												
Lys	Lys	Tyr	Arg	Asn	Met	Val	Val	Arg	Ala	Cys										
				90				95												
Gly	Cys	His																		
100																				

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 106 amino acids

(B) TYPE: protein

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: human

(F) TISSUE TYPE: BRAIN

## (ix) FEATURE:

(D) OTHER INFORMATION:

/product= "GDF-1 (fx)"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Cys Arg Ala Arg Arg Leu Tyr Val Ser Phe Arg Glu Val Gly  
 1 5 10  
 Trp His Arg Trp Val Ile Ala Pro Arg Gly Phe Leu Ala Asn Tyr  
 15 20 25  
 Cys Gln Gly Gln Cys Ala Leu Pro Val Ala Leu Ser Gly Ser Gly  
 30 35 40  
 Gly Pro Pro Ala Leu Asn His Ala Val Leu Arg Ala Leu Met His  
 45 50 55  
 Ala Ala Ala Pro Gly Ala Ala Asp Leu Pro Cys Cys Val Pro Ala  
 60 65 70  
 Arg Leu Ser Pro Ile Ser Val Leu Phe Phe Asp Asn Ser Asp Asn  
 75 80 85  
 Val Val Leu Arg Gln Tyr Glu Asp Met Val Val Asp Glu Cys Gly  
 90 95 100  
 Cys Arg  
 105

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Cys Xaa Xaa Xaa Xaa  
 1 5

## (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1822 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: HOMO SAPIENS
  - (F) TISSUE TYPE: HIPPOCAMPUS
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 49..1341
  - (D) OTHER INFORMATION: /standard\_name= "hOP1"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

GGTGCGGGCC CGGAGCCCGG AGCCCGGGTA GCGCGTAGAG CCGGCGCG ATG CAC GTG      57
                                         Met His Val
                                         1

CGC TCA CTG CGA GCT GCG GCG CCG CAC AGC TTC GTG GCG CTC TGG GCA      105
Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala
   5              10              15

CCC CTG TTC CTG CTG CGC TCC GCC CTG GCC GAC TTC AGC CTG GAC AAC      153
Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn
  20              25              30              35

GAG GTG CAC TCG AGC TTC ATC CAC CGG CGC CTC CGC AGC CAG GAG CGG      201
Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg
              40              45              50

CGG GAG ATG CAG CGC GAG ATC CTC TCC ATT TTG GGC TTG CCC CAC CGC      249
Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg
              55              60              65

CCG CGC CCG CAC CTC CAG GGC AAG CAC AAC TCG GCA CCC ATG TTC ATG      297
Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met
              70              75              80

CTG GAC CTG TAC AAC GCC ATG GCG GTG GAG GAG GGC GGC GGG CCC GGC      345
Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly Gly Pro Gly
   85              90              95

GGC CAG GGC TTC TCC TAC CCC TAC AAG GCC GTC TTC AGT ACC CAG GGC      393
Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly
100              105              110              115

```

CCC CCT CTG GCC AGC CTG CAA GAT AGC CAT TTC CTC ACC GAC GCC GAC Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp 120 125 130	441
ATG GTC ATG AGC TTC GTC AAC CTC GTG GAA CAT GAC AAG GAA TTC TTC Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe 135 140 145	489
CAC CCA CGC TAC CAC CAT CGA GAG TTC CGG TTT GAT CTT TCC AAG ATC His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile 150 155 160	537
CCA GAA GGG GAA GCT GTC ACG GCA GCC GAA TTC CGG ATC TAC AAG GAC Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp 165 170 175	585
TAC ATC CGG GAA CGC TTC GAC AAT GAG ACG TTC CGG ATC AGC GTT TAT Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile Ser Val Tyr 180 185 190 195	633
CAG GTG CTC CAG GAG CAC TTG GGC AGG GAA TCG GAT CTC TTC CTG CTC Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu Phe Leu Leu 200 205 210	681
GAC AGC CGT ACC CTC TGG GCC TCG GAG GAG GGC TGG CTG GTG TTT GAC Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp 215 220 225	729
ATC ACA GCC ACC AGC AAC CAC TGG GTG GTC AAT CCG CGG CAC AAC CTG Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu 230 235 240	777
GGC CTG CAG CTC TCG GTG GAG ACG CTG GAT GGG CAG AGC ATC AAC CCC Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro 245 250 255	825
AAG TTG GCG GGC CTG ATT GGG CGG CAC GGG CCC CAG AAC AAG CAG CCC Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro 260 265 270 275	873
TTC ATG GTG GCT TTC TTC AAG GCC ACG GAG GTC CAC TTC CGC AGC ATC Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe Arg Ser Ile 280 285 290	921
CGG TCC ACG GGG AGC AAA CAG CGC AGC CAG AAC CGC TCC AAG ACG CCC Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro 295 300 305	969
AAG AAC CAG GAA GCC CTG CGG ATG GCC AAC GTG GCA GAG AAC AGC AGC Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser 310 315 320	1017



AGC GAC CAG AGG CAG GCC TGT AAG AAG CAC GAG CTG TAT GTC AGC TTC Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe 325 330 335	1065
CGA GAC CTG GGC TGG CAG GAC TGG ATC ATC GCG CCT GAA GGC TAC GCC Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala 340 345 350 355	1113
GCC TAC TAC TGT GAG GGG GAG TGT GCC TTC CCT CTG AAC TCC TAC ATG Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met 360 365 370	1161
AAC GCC ACC AAC CAC GCC ATC GTG CAG ACG CTG GTC CAC TTC ATC AAC Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn 375 380 385	1209
CCG GAA ACG GTG CCC AAG CCC TGC TGT GCG CCC ACG CAG CTC AAT GCC Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala 390 395 400	1257
ATC TCC GTC CTC TAC TTC GAT GAC AGC TCC AAC GTC ATC CTG AAG AAA Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys 405 410 415	1305
TAC AGA AAC ATG GTG GTC CGG GCC TGT GGC TGC CAC TAGCTCCTCC Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His 420 425 430	1351
GAGAATTCAG ACCCTTTGGG GCCAAGTTTT TCTGGATCCT CCATTGCTCG CCTTGGCCAG	1411
GAACCAGCAG ACCAACTGCC TTTTGTGAGA CTTTCCCCTC CCTATCCCCA ACTTTAAAGG	1471
TGTGAGAGTA TTAGGAAACA TGAGCAGCAT ATGGCTTTTG ATCAGTTTTT CAGTGGCAGC	1531
ATCCAATGAA CAAGATCCTA CAAGCTGTGC AGGCAAAACC TAGCAGGAAA AAAAAACAAC	1591
GCATAAAGAA AAATGGCCCG GCCAGGTCAT TGGCTGGGAA GTCTCAGCCA TGCACGGACT	1651
CGTTTCCAGA GGTAATTATG AGCGCCTACC AGCCAGGCCA CCCAGCCGTG GGAGGAAGGG	1711
GGCGTGGCAA GGGGTGGGCA CATTGGTGTC TGTGCGAAAG GAAAATTGAC CCGGAAGTTC	1771
CTGTAATAAA TGTACAATA AAACGAATGA ATGAAAAAAA AAAAAAAAAA A	1822

## (2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 431 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(D) OTHER INFORMATION: /Product="OP1-PP"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
 1           5           10           15
Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
          20           25           30
Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
          35           40           45
Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
          50           55           60
Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
          65           70           75           80
Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly
          85           90           95
Gly Pro Gly Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser
          100          105          110
Thr Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr
          115          120          125
Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys
          130          135          140
Glu Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu
          145          150          155          160
Ser Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile
          165          170          175
Tyr Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile
          180          185          190
Ser Val Tyr Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu
          195          200          205
Phe Leu Leu Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu
          210          215          220
Val Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg
          225          230          235          240
His Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser
          245          250          255

```

```

Ile Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn
      260                      265                      270
Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe
      275                      280                      285
Arg Ser Ile Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser
      290                      295                      300
Lys Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu
      305                      310                      315                      320
Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr
      325                      330                      335
Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu
      340                      345                      350
Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn
      355                      360                      365
Ser Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His
      370                      375                      380
Phe Ile Asn Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln
      385                      390                      395                      400
Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile
      405                      410                      415
Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
      420                      425                      430

```

## (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1873 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: MURIDAE
  - (F) TISSUE TYPE: EMBRYO
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 104..1393
  - (D) OTHER INFORMATION: /note= "MOP1 (CDNA)"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTGCAGCAAG TGACCTCGGG TCGTGGACCG CTGCCCTGCC CCCTCCGCTG CCACCTGGGG	60
CGGCGCGGGC CCGGTGCCCC GGATCGCGCG TAGAGCCGGC GCG ATG CAC GTG CGC	115
Met His Val Arg	1
TCG CTG CGC GCT GCG GCG CCA CAC AGC TTC GTG GCG CTC TGG GCG CCT	163
Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala Pro	5 10 15 20
CTG TTC TTG CTG CGC TCC GCC CTG GCC GAT TTC AGC CTG GAC AAC GAG	211
Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn Glu	25 30 35
GTG CAC TCC AGC TTC ATC CAC CGG CGC CTC CGC AGC CAG GAG CGG CGG	259
Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg Arg	40 45 50
GAG ATG CAG CGG GAG ATC CTG TCC ATC TTA GGG TTG CCC CAT CGC CCG	307
Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg Pro	55 60 65
CGC CCG CAC CTC CAG GGA AAG CAT AAT TCG GCG CCC ATG TTC ATG TTG	355
Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met Leu	70 75 80
GAC CTG TAC AAC GCC ATG GCG GTG GAG GAG AGC GGG CCG GAC GGA CAG	403
Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Ser Gly Pro Asp Gly Gln	85 90 95 100
GGC TTC TCC TAC CCC TAC AAG GCC GTC TTC AGT ACC CAG GGC CCC CCT	451
Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly Pro Pro	105 110 115
TTA GCC AGC CTG CAG GAC AGC CAT TTC CTC ACT GAC GCC GAC ATG GTC	499
Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp Met Val	120 125 130
ATG AGC TTC GTC AAC CTA GTG GAA CAT GAC AAA GAA TTC TTC CAC CCT	547
Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe His Pro	135 140 145
CGA TAC CAC CAT CGG GAG TTC CGG TTT GAT CTT TCC AAG ATC CCC GAG	595
Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile Pro Glu	150 155 160
GGC GAA GCG GTG ACC GCA GCC GAA TTC AGG ATC TAT AAG GAC TAC ATC	643
Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp Tyr Ile	165 170 175 180

CGG GAG CGA TTT GAC AAC GAG ACC TTC CAG ATC ACA GTC TAT CAG TGG Arg Glu Arg Phe Asp Asn Glu Thr Phe Gln Ile Thr Val Tyr Gln Trp 185 190 195	691
CTC CAG GAG CAC TCA GGC AGG GAG TCG GAC CTC TTC TTG CTG GAC AGC Leu Gln Glu His Ser Gly Arg Glu Ser Asp Leu Phe Leu Leu Asp Ser 200 205 210	739
CGC ACC ATC TGG GCT TCT GAG GAG GGC TGG TTG GTG TTT GAT ATC ACA Arg Thr Ile Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp Ile Thr 215 220 225	787
GCC ACC AGC AAC CAC TGG GTG GTC AAC CCT CGG CAC AAC CTG GGC TTA Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu Gly Leu 230 235 240	835
CAG CTC TCT GTG GAG ACC CTG GAT GGG CAG AGC ATC AAC CCC AAG TTG Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro Lys Leu 245 250 255 260	883
GCA GGC CTG ATT GGA CGG CAT GGA CCC CAG AAC AAG CAA CCC TTC ATG Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro Phe Met 265 270 275	931
GTG GCC TTC TTC AAG GCC ACG GAA GTC CAT CTC CGT AGT ATC CGG TCC Val Ala Phe Phe Lys Ala Thr Glu Val His Leu Arg Ser Ile Arg Ser 280 285 290	979
ACG GGG GGC AAG CAG CGC AGC CAG AAT CGC TCC AAG ACG CCA AAG AAC Thr Gly Gly Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro Lys Asn 295 300 305	1027
CAA GAG GCC CTG AGG ATG GCC AGT GTG GCA GAA AAC AGC AGC AGT GAC Gln Glu Ala Leu Arg Met Ala Ser Val Ala Glu Asn Ser Ser Ser Asp 310 315 320	1075
CAG AGG CAG GCC TGC AAG AAA CAT GAG CTG TAC GTC AGC TTC CGA GAC Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp 325 330 335 340	1123
CTT GGC TGG CAG GAC TGG ATC ATT GCA CCT GAA GGC TAT GCT GCC TAC Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr 345 350 355	1171
TAC TGT GAG GGA GAG TGC GCC TTC CCT CTG AAC TCC TAC ATG AAC GCC Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala 360 365 370	1219
ACC AAC CAC GCC ATC GTC CAG ACA CTG GTT CAC TTC ATC AAC CCA GAC Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro Asp 375 380 385	1267

Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala  
1 5 10 15

Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser  
20 25 30

Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser  
35 40 45

Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu  
 50 55 60  
 Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro  
 65 70 75 80  
 Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Ser Gly  
 85 90 95  
 Pro Asp Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr  
 100 105 110  
 Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp  
 115 120 125  
 Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu  
 130 135 140  
 Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser  
 145 150 155 160  
 Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr  
 165 170 175  
 Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Gln Ile Thr  
 180 185 190  
 Val Tyr Gln Trp Leu Gln Glu His Ser Gly Arg Glu Ser Asp Leu Phe  
 195 200 205  
 Leu Leu Asp Ser Arg Thr Ile Trp Ala Ser Glu Glu Gly Trp Leu Val  
 210 215 220  
 Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His  
 225 230 235 240  
 Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile  
 245 250 255  
 Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys  
 260 265 270  
 Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Leu Arg  
 275 280 285  
 Ser Ile Arg Ser Thr Gly Gly Lys Gln Arg Ser Gln Asn Arg Ser Lys  
 290 295 300  
 Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Ser Val Ala Glu Asn  
 305 310 315 320  
 Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val  
 325 330 335

Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly  
                   340                  345                  350

Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser  
                   355                  360                  365

Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe  
                   370                  375                  380

Ile Asn Pro Asp Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu  
                   385                  390                  395                  400

Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu  
                   405                  410                  415

Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His  
                   420                  425                  430

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1723 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (F) TISSUE TYPE: HIPPOCAMPUS

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 490..1696
- (D) OTHER INFORMATION: /note= "hOP2 (cDNA)"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGCGCCGGCA GAGCAGGAGT GGCTGGAGGA GCTGTGGTGTG GAGCAGGAGG TGGCACGGCA	60
GGGCTGGAGG GCTCCCTATG AGTGGCGGAG ACGGCCCAGG AGGCGCTGGA GCAACAGCTC	120
CCACACCGCA CCAAGCGGTG GCTGCAGGAG CTCGCCCATC GCCCCTGCGC TGCTCGGACC	180
GCGGCCACAG CCGGACTGGC GGGTACGGCG GCGACAGAGG CATTGGCCGA GAGTCCCAGT	240
CCGCAGAGTA GCCCCGGCCT CGAGGCGGTG GCGTCCCGGT CCTCTCCGTC CAGGAGCCAG	300
GACAGGTGTC GCGCGGCGGG GCTCCAGGGA CCGCGCCTGA GGCCGGCTGC CCGCCCGTCC	360
CGCCCCGCCC CGCCGCCCGC CGCCGCCCGA GCCCAGCCTC CTTGCCGTCTG GGGCGTCCCC	420



AGGCCCTGGG TCGGCCGCGG AGCCGATGCG CGCCCGCTGA GCGCCCCAGC TGAGCGCCCC	480
CGGCCCTGCC ATG ACC GCG CTC CCC GGC CCG CTC TGG CTC CTG GGC CTG Met Thr Ala Leu Pro Gly Pro Leu Trp Leu Leu Gly Leu 1 5 10	528
GCG CTA TGC GCG CTG GGC GGG GGC GGC CCC GGC CTG CGA CCC CCG CCC Ala Leu Cys Ala Leu Gly Gly Gly Gly Pro Gly Leu Arg Pro Pro Pro 15 20 25	576
GGC TGT CCC CAG CGA CGT CTG GGC GCG CGC GAG CGC CGG GAC GTG CAG Gly Cys Pro Gln Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Val Gln 30 35 40 45	624
CGC GAG ATC CTG GCG GTG CTC GGG CTG CCT GGG CGG CCC CGG CCC CGC Arg Glu Ile Leu Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg 50 55 60	672
GCG CCA CCC GCC GCC TCC CGG CTG CCC GCG TCC GCG CCG CTC TTC ATG Ala Pro Pro Ala Ala Ser Arg Leu Pro Ala Ser Ala Pro Leu Phe Met 65 70 75	720
CTG GAC CTG TAC CAC GCC ATG GCC GGC GAC GAC GAC GAG GAC GGC GCG Leu Asp Leu Tyr His Ala Met Ala Gly Asp Asp Asp Glu Asp Gly Ala 80 85 90	768
CCC GCG GAG CGG CGC CTG GGC CGC GCC GAC CTG GTC ATG AGC TTC GTT Pro Ala Glu Arg Arg Leu Gly Arg Ala Asp Leu Val Met Ser Phe Val 95 100 105	816
AAC ATG GTG GAG CGA GAC CGT GCC CTG GGC CAC CAG GAG CCC CAT TGG Asn Met Val Glu Arg Asp Arg Ala Leu Gly His Gln Glu Pro His Trp 110 115 120 125	864
AAG GAG TTC CGC TTT GAC CTG ACC CAG ATC CCG GCT GGG GAG GCG GTC Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val 130 135 140	912
ACA GCT GCG GAG TTC CGG ATT TAC AAG GTG CCC AGC ATC CAC CTG CTC Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val Pro Ser Ile His Leu Leu 145 150 155	960
AAC AGG ACC CTC CAC GTC AGC ATG TTC CAG GTG GTC CAG GAG CAG TCC Asn Arg Thr Leu His Val Ser Met Phe Gln Val Val Gln Glu Gln Ser 160 165 170	1008
AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT CTT CAG ACG CTC CGA GCT Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln Thr Leu Arg Ala 175 180 185	1056
GGA GAC GAG GGC TGG CTG GTG CTG GAT GTC ACA GCA GCC AGT GAC TGC Gly Asp Glu Gly Trp Leu Val Leu Asp Val Thr Ala Ala Ser Asp Cys 190 195 200 205	1104

TGG TTG CTG AAG CGT CAC AAG GAC CTG GGA CTC CGC CTC TAT GTG GAG Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu 210 215 220	1152
ACT GAG GAC GGG CAC AGC GTG GAT CCT GGC CTG GCC GGC CTG CTG GGT Thr Glu Asp Gly His Ser Val Asp Pro Gly Leu Ala Gly Leu Leu Gly 225 230 235	1200
CAA CGG GCC CCA CGC TCC CAA CAG CCT TTC GTG GTC ACT TTC TTC AGG Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val Thr Phe Phe Arg 240 245 250	1248
GCC AGT CCG AGT CCC ATC CGC ACC CCT CGG GCA GTG AGG CCA CTG AGG Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg Ala Val Arg Pro Leu Arg 255 260 265	1296
AGG AGG CAG CCG AAG AAA AGC AAC GAG CTG CCG CAG GCC AAC CGA CTC Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu Pro Gln Ala Asn Arg Leu 270 275 280 285	1344
CCA GGG ATC TTT GAT GAC GTC CAC GGC TCC CAC GGC CGG CAG GTC TGC Pro Gly Ile Phe Asp Asp Val His Gly Ser His Gly Arg Gln Val Cys 290 295 300	1392
CGT CGG CAC GAG CTC TAC GTC AGC TTC CAG GAC CTC GGC TGG CTG GAC Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu Asp 305 310 315	1440
TGG GTC ATC GCT CCC CAA GGC TAC TCG GCC TAT TAC TGT GAG GGG GAG Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu 320 325 330	1488
TGC TCC TTC CCA CTG GAC TCC TGC ATG AAT GCC ACC AAC CAC GCC ATC Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala Ile 335 340 345	1536
CTG CAG TCC CTG GTG CAC CTG ATG AAG CCA AAC GCA GTC CCC AAG GCG Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala Val Pro Lys Ala 350 355 360 365	1584
TGC TGT GCA CCC ACC AAG CTG AGC GCC ACC TCT GTG CTC TAC TAT GAC Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp 370 375 380	1632
AGC AGC AAC AAC GTC ATC CTG CGC AAA CAC CGC AAC ATG GTG GTC AAG Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn Met Val Val Lys 385 390 395	1680
GCC TGC GGC TGC CAC T GAGTCAGCCC GCCCAGCCCT ACTGCAG Ala Cys Gly Cys His 400	1723

## (2) INFORMATION FOR SEQ ID NO:21:

## (i)SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 402 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii)MOLECULE TYPE: protein

## (ix)FEATURE:

(A)OTHER INFORMATION: /product= "hOP2-PP"

## (xi)SEQUENCE DESCRIPTION: SEQ ID NO:21:

```

Met Thr Ala Leu Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu Cys
 1               5               10               15
Ala Leu Gly Gly Gly Gly Pro Gly Leu Arg Pro Pro Pro Gly Cys Pro
      20               25               30
Gln Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Val Gln Arg Glu Ile
      35               40               45
Leu Ala Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg Ala Pro Pro
      50               55               60
Ala Ala Ser Arg Leu Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu
      65               70               75               80
Tyr His Ala Met Ala Gly Asp Asp Asp Glu Asp Gly Ala Pro Ala Glu
      85               90               95
Arg Arg Leu Gly Arg Ala Asp Leu Val Met Ser Phe Val Asn Met Val
      100              105              110
Glu Arg Asp Arg Ala Leu Gly His Gln Glu Pro His Trp Lys Glu Phe
      115              120              125
Arg Phe Asp Leu Thr Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala
      130              135              140
Glu Phe Arg Ile Tyr Lys Val Pro Ser Ile His Leu Leu Asn Arg Thr
      145              150              155              160
Leu His Val Ser Met Phe Gln Val Val Gln Glu Gln Ser Asn Arg Glu
      165              170              175
Ser Asp Leu Phe Phe Leu Asp Leu Gln Thr Leu Arg Ala Gly Asp Glu
      180              185              190
Gly Trp Leu Val Leu Asp Val Thr Ala Ala Ser Asp Cys Trp Leu Leu
      195              200              205

```

Lys Arg His Lys Asp Leu Gly Leu Arg Leu Tyr Val Glu Thr Glu Asp  
 210 215 220  
 Gly His Ser Val Asp Pro Gly Leu Ala Gly Leu Leu Gly Gln Arg Ala  
 225 230 235 240  
 Pro Arg Ser Gln Gln Pro Phe Val Val Thr Phe Phe Arg Ala Ser Pro  
 245 250 255  
 Ser Pro Ile Arg Thr Pro Arg Ala Val Arg Pro Leu Arg Arg Arg Gln  
 260 265 270  
 Pro Lys Lys Ser Asn Glu Leu Pro Gln Ala Asn Arg Leu Pro Gly Ile  
 275 280 285  
 Phe Asp Asp Val His Gly Ser His Gly Arg Gln Val Cys Arg Arg His  
 290 295 300  
 Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu Asp Trp Val Ile  
 305 310 315 320  
 Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly Glu Cys Ser Phe  
 325 330 335  
 Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala Ile Leu Gln Ser  
 340 345 350  
 Leu Val His Leu Met Lys Pro Asn Ala Val Pro Lys Ala Cys Cys Ala  
 355 360 365  
 Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn  
 370 375 380  
 Asn Val Ile Leu Arg Lys His Arg Asn Met Val Val Lys Ala Cys Gly  
 385 390 395 400  
 Cys His

## (2) INFORMATION FOR SEQ ID NO:22:

- (1) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1926 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: MURIDAE
  - (F) TISSUE TYPE: EMBRYO
- (ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 93..1289  
 (D) OTHER INFORMATION: /note= "mOP2 cDNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCCAGGCACA GGTGCGCCGT CTGGTCCTCC CCGTCTGGCG TCAGCCGAGC	50
CCGACCAGCT ACCAGTGGAT GCGCGCCGGC TGAAAGTCCG AG ATG GCT ATG CGT	104
Met Ala Met Arg	
1	
CCC GGG CCA CTC TGG CTA TTG GGC CTT GCT CTG TGC GCG CTG GGA GGC	152
Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu Cys Ala Leu Gly Gly	
5 10 15 20	
GGC CAC GGT CCG CGT CCC CCG CAC ACC TGT CCC CAG CGT CGC CTG GGA	200
Gly His Gly Pro Arg Pro Pro His Thr Cys Pro Gln Arg Arg Leu Gly	
25 30 35	
GCG CGC GAG CGC CGC GAC ATG CAG CGT GAA ATC CTG GCG GTG CTC GGG	248
Ala Arg Glu Arg Arg Asp Met Gln Arg Glu Ile Leu Pro Val Leu Gly	
40 45 50	
CTA CCG GGA CGG CCC CGA CCC CGT GCA CAA CCC GCG GCT GCC CGG CAG	296
Leu Pro Gly Arg Pro Arg Pro Arg Ala Gln Pro Ala Ala Ala Arg Gln	
55 60 65	
CCA GCG TCC GCG CCC CTC TTC ATG TTG GAC CTA TAC CAC GCC ATG ACC	344
Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr His Ala Met Thr	
70 75 80	
GAT GAC GAC GAC GGC GGG CCA CCA CAG GCT CAC TTA GGC CGT GCC GAC	392
Asp Asp Asp Asp Gly Gly Pro Pro Gln Ala His Leu Gly Arg Ala Asp	
85 90 95 100	
CTG GTC ATG AGC TTC GTC AAC ATG GTG GAA CGC GAC CGT ACC CTG GGC	440
Leu Val Met Ser Phe Val Asn Met Val Glu Arg Asp Arg Thr Leu Gly	
105 110 115	
TAC CAG GAG CCA CAC TGG AAG GAA TTC CAC TTT GAC CTA ACC CAG ATC	488
Tyr Gln Glu Pro His Trp Lys Glu Phe His Phe Asp Leu Thr Gln Ile	
120 125 130	
CCT GCT GGG GAG GCT GTC ACA GCT GCT GAG TTC CGG ATC TAC AAA GAA	536
Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Glu	
135 140 145	
CCC AGC ACC CAC CCG CTC AAC ACA ACC CTC CAC ATC AGC ATG TTC GAA	584
Pro Ser Thr His Pro Leu Asn Thr Thr Leu His Ile Ser Met Phe Glu	
150 155 160	

GTG GTC CAA GAG CAC TCC AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT Val Val Gln Glu His Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp 165 170 175 180	632
CTT CAG ACG CTC CGA TCT GGG GAC GAG GGC TGG CTG GTG CTG GAC ATC Leu Gln Thr Leu Arg Ser Gly Asp Glu Gly Trp Leu Val Leu Asp Ile 185 190 195	680
ACA GCA GCC AGT GAC CGA TGG CTG CTG AAC CAT CAC AAG GAC CTG GGA Thr Ala Ala Ser Asp Arg Trp Leu Leu Asn His His Lys Asp Leu Gly 200 205 210	728
CTC CGC CTC TAT GTG GAA ACC GCG GAT GGG CAC AGC ATG GAT CCT GGC Leu Arg Leu Tyr Val Glu Thr Ala Asp Gly His Ser Met Asp Pro Gly 215 220 225	776
CTG GCT GGT CTG CTT GGA CGA CAA GCA CCA CGC TCC AGA CAG CCT TTC Leu Ala Gly Leu Leu Gly Arg Gln Ala Pro Arg Ser Arg Gln Pro Phe 230 235 240	824
ATG GTA ACC TTC TTC AGG GCC AGC CAG AGT CCT GTG CGG GCC CCT CGG Met Val Thr Phe Phe Arg Ala Ser Gln Ser Pro Val Arg Ala Pro Arg 245 250 255 260	872
GCA GCG AGA CCA CTG AAG AGG AGG CAG CCA AAG AAA ACG AAC GAG CTT Ala Ala Arg Pro Leu Lys Arg Arg Gln Pro Lys Lys Thr Asn Glu Leu 265 270 275	920
CCG CAC CCC AAC AAA CTC CCA GGG ATC TTT GAT GAT GGC CAC GGT TCC Pro His Pro Asn Lys Leu Pro Gly Ile Phe Asp Asp Gly His Gly Ser 280 285 290	968
CGC GGC AGA GAG GTT TGC CGC AGG CAT GAG CTC TAC GTC AGC TTC CGT Arg Gly Arg Glu Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Arg 295 300 305	1016
GAC CTT GGC TGG CTG GAC TGG GTC ATC GCC CCC CAG GGC TAC TCT GCC Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala 310 315 320	1064
TAT TAC TGT GAG GGG GAG TGT GCT TTC CCA CTG GAC TCC TGT ATG AAC Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asp Ser Cys Met Asn 325 330 335 340	1112
GCC ACC AAC CAT GCC ATC TTG CAG TCT CTG GTG CAC CTG ATG AAG CCA Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro 345 350 355	1160
GAT GTT GTC CCC AAG GCA TGC TGT GCA CCC ACC AAA CTG AGT GCC ACC Asp Val Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr 360 365 370	1208

```

TCT GTG CTG TAC TAT GAC AGC AGC AAC AAT GTC ATC CTG CGT AAA CAC      1256
Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His
      375                      380                      385

CGT AAC ATG GTG GTC AAG GCC TGT GGC TGC CAC TGAGGCCCGG CCCAGCATCC      1309
Arg Asn Met Val Val Lys Ala Cys Gly Cys His
      390                      395

TGCTTCTACT ACCTTACCAT CTGGCCGGGC CCCTCTCCAG AGGCAGAAAC CCTTCTATGT      1369

TATCATAGCT CAGACAGGGG CAATGGGAGG CCCTTCACTT CCCCTGGCCA CTTCCTGCTA      1429

AAATTCTGGT CTTTCCCAGT TCCTCTGTCC TTCATGGGGT TTCGGGGCTA TCACCCCGCC      1489

CTCTCCATCC TCCTACCCCA AGCATAGACT GAATGCACAC AGCATCCCAG AGCTATGCTA      1549

ACTGAGAGGT CTGGGGTCAG CACTGAAGGC CCACATGAGG AAGACTGATC CTTGGCCATC      1609

CTCAGCCCAC AATGGCAAAT TCTGGATGGT CTAAGAAGGC CGTGGAATTC TAAACTAGAT      1669

GATCTGGGCT CTCTGCACCA TTCATTGTGG CAGTTGGGAC ATTTTGTAGGT ATAACAGACA      1729

CATACACTTA GATCAATGCA TCGCTGTACT CCTTGAAATC AGAGCTAGCT TGTTAGAAAA      1789

AGAAATCAGAG CCAGGTATAG CGGTGCATGT CATTAAATCCC AGCGCTAAAG AGACAGAGAC      1849

AGGAGAATCT CTGTGAGTTC AAGGCCACAT AGAAAGAGCC TGTCTCGGGA GCAGGAAAAA      1909

AAAAAAAAAC GGAATTC                                                    1926

```

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (ix) FEATURE:

- (D) OTHER INFORMATION: /product= "mOP2-PP"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```

Met Ala Met Arg Pro Gly Pro Leu Trp Leu Leu Gly Leu Ala Leu Cys
 1           5           10           15

Ala Leu Gly Gly Gly His Gly Pro Arg Pro Pro His Thr Cys Pro Gln
      20           25           30

Arg Arg Leu Gly Ala Arg Glu Arg Arg Asp Met Gln Arg Glu Ile Leu Ala
 35           40           45

```

Val Leu Gly Leu Pro Gly Arg Pro Arg Pro Arg Ala Gln Pro Ala Ala  
 50 55 60 65  
 Ala Arg Gln Pro Ala Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr His Ala  
 70 75 80  
 Met Thr Asp Asp Asp Asp Gly Gly Pro Pro Gln Ala His Leu Gly Arg  
 85 90 95  
 Ala Asp Leu Val Met Ser Phe Val Asn Met Val Glu Arg Asp Arg Thr  
 100 105 110  
 Leu Gly Tyr Gln Glu Pro His Trp Lys Glu Phe His Phe Asp Leu Thr  
 115 120 125 130  
 Gln Ile Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr  
 135 140 145  
 Lys Glu Pro Ser Thr His Pro Leu Asn Thr Thr Leu His Ile Ser Met  
 150 155 160  
 Phe Glu Val Val Gln Glu His Ser Asn Arg Glu Ser Asp Leu Phe Phe  
 165 170 175  
 Leu Asp Leu Gln Thr Leu Arg Ser Gly Asp Glu Gly Trp Leu Val Leu  
 180 185 190  
 Asp Ile Thr Ala Ala Ser Asp Arg Trp Leu Leu Asn His His Lys Asp  
 195 200 205 210  
 Leu Gly Leu Arg Leu Tyr Val Glu Thr Ala Asp Gly His Ser Met Asp  
 215 220 225  
 Pro Gly Leu Ala Gly Leu Leu Gly Arg Gln Ala Pro Arg Ser Arg Gln  
 230 235 240  
 Pro Phe Met Val Thr Phe Phe Arg Ala Ser Gln Ser Pro Val Arg Ala  
 245 250 255  
 Pro Arg Ala Ala Arg Pro Leu Lys Arg Arg Gln Pro Lys Lys Thr Asn  
 260 265 270  
 Glu Leu Pro His Pro Asn Lys Leu Pro Gly Ile Phe Asp Asp Gly His  
 275 280 285 290  
 Gly Ser Arg Gly Arg Glu Val Cys Arg Arg His Glu Leu Tyr Val Ser  
 295 300 305  
 Phe Arg Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr  
 310 315 320



Ser Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asp Ser Cys  
325 330 335

Met Asn Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met  
340 345 350

Lys Pro Asp Val Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser  
355 360 365 370

Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg  
375 380 385

Lys His Arg Asn Met Val Val Lys Ala Cys Gly Cys His  
390 395

What is claimed is:

1. A composition for increasing the progenitor cell population in a mammal comprising:  
5 progenitor cells, stimulated ex vivo by exposure to a morphogen at a concentration and for a time sufficient such that said progenitor cells are stimulated to proliferate.
- 10 2. A composition for inducing non-chondrogenic tissue growth in a mammal comprising:  
progenitor cells, stimulated by exposure to a morphogen at a concentration and for a time sufficient such that said progenitor cells, when  
15 disposed in vivo within a tissue locus, are capable of non-chondrogenic tissue-specific differentiation and proliferation within said locus.
3. The composition of claim 1 or 2 wherein  
20 said progenitor cells are hemopoietic pluripotential stem cells.
4. The composition of claim 1 or 2 wherein  
said progenitor cells are of mesenchymal origin.  
25
5. A composition for inducing the formation of non-chondrogenic replacement tissue at a tissue locus in a mammal comprising:  
a biocompatible, acellular matrix  
30 having components specific for said tissue and capable of providing a morphogenically permissive, tissue-specific environment; and  
a morphogen such that said morphogen, when absorbed on said matrix and provided to a

tissue-specific locus requiring replacement tissue,  
is capable of inducing the developmental cascade of  
tissue morphogenesis at said locus.

- 5 6. A composition for inducing the  
formation of non-chondrogenic replacement tissue at a  
tissue locus in a mammal comprising:  
a biocompatible, acellular matrix  
capable of providing a morphogenically permissive  
10 environment; and  
a morphogen such that said morphogen,  
when absorbed on said matrix and provided to a  
tissue-specific locus requiring replacement tissue,  
is capable of inducing the developmental cascade of  
15 tissue morphogenesis at said locus.
7. The composition of claim 5 or 6 wherein  
said matrix is biodegradable.
- 20 8. The composition of claim 5 or 6 wherein  
said matrix is derived from organ-specific tissue.
9. The composition of claim 5 or 6 wherein  
said matrix comprises collagen and cell attachment  
25 factors selected from the group consisting of  
glycosaminoglycans and proteoglycans.
10. The composition of claim 5 or 6 wherein  
said matrix defines pores of a dimension sufficient  
30 to permit the influx, differentiation and  
proliferation of migratory progenitor cells from the  
body of said mammal.

11. The composition of claim 1, 2, 5 or 6 wherein said morphogen comprises an amino acid sequence sharing at least 70% homology with one of the sequences selected from the group consisting of:
- 5 hOP1 (Seq. ID No. 5); mOP1 (Seq. ID No. 6); hOP2 (Seq. ID No. 7); mOP2 (Seq. ID No. 8); CBMP2A(fx) (Seq. ID No. 9); CBMP2B(fx) (Seq. ID No. 10); DPP(fx) (Seq. ID No. 11); Vgl(fx) (Seq. ID No. 12); Vgr-1(fx) (Seq. ID No. 13); and GDF-1(fx) (Seq. ID No. 14).
- 10
12. The composition of claim 11 wherein said morphogen comprises an amino acid sequence sharing at least 80% homology with one of the sequences selected from said group.
- 15
13. The composition of claim 12 wherein said morphogen comprises an amino acid sequence having greater than 60% amino acid identity with the sequence defined by residues 43-139 of Seq. ID No.5
- 20 (hOP1).
14. The composition of claim 13 wherein said morphogen comprises an amino acid sequence having greater than 65% identity with the sequence defined
- 25 by residues 43-139 of Seq. ID No.5 (hOP1).
15. A method of increasing a population of progenitor cells comprising the step of:
- contacting progenitor cells with a
- 30 morphogen at a concentration and for a time sufficient such that said progenitor cells are stimulated to proliferate.

16. The method of claim 15 for increasing progenitor cells in a mammal comprising the additional step of supplying said stimulated progenitor cells to a mammal to increase the  
5 progenitor cell population in said mammal.

17. A method of inducing non-chondrogenic tissue growth in a mammal comprising the step of:  
contacting progenitor cells with a  
10 morphogen at a concentration and for a time sufficient such that said progenitor cells, when provided to a tissue-specific locus in a mammal, are capable of nonchondrogenic tissue-specific differentiation and proliferation at said locus.

15 18. The method of claim 14 or 16 wherein said progenitor cells are of mesenchymal origin.

19. A method of maintaining the phenotypic  
20 expression of differentiated cells in a mammal comprising the steps of:  
contacting said differentiated cells with a morphogen at a concentration and for a time sufficient such that said cells are stimulated to  
25 express their phenotype.

20. The method of claim 19 wherein said differentiated cells are senescent or quiescent cells.

30 21. A method of inducing non-chondrogenic tissue growth at a tissue locus in a mammal comprising:

providing said locus with a morphogen  
at a concentration and for a time sufficient such  
that said protein, when provided to a morphogenically  
permissive tissue-specific locus, is capable of  
5 inducing the developmental cascade of tissue  
morphogenesis at said locus.

22. The method of claim 21 wherein said  
nonchondrogenic tissue is hepatic tissue, and said  
10 tissue locus is the liver.

23. The method of claim 22 wherein said  
protein is provided to said locus in association with  
a biocompatible, acellular matrix.  
15

24. The method of claim 23 wherein said  
matrix has components specific for said tissue.

25. The method of claim 23 wherein said  
20 matrix is biodegradable.

26. The method of claim 23 wherein said  
matrix is derived from organ-specific tissue.

25 27. The method of claim 23 wherein said  
matrix comprises collagen and cell attachment factors  
specific for said tissue.

28. The method of claim 23 wherein said  
30 matrix defines pores of a dimension sufficient to  
permit the influx, differentiation and proliferation  
of migratory progenitor cells from the body of said  
mammal.

29. The method of claim 14, 16, 17 or 20 where said morphogen comprises an amino acid sequence sharing at least 70% homology with one of the sequences selected from the group consisting of hOP1  
5 (Seq. ID No. 5); mOP1 (Seq. ID No. 6); hOP2 (Seq. ID No. 7); mOP2 (Seq. ID No. 8); CBMP2A(fx) (Seq. ID No. 9); CBMP2B(fx) (Seq. ID No. 10); DPP(fx) (Seq. ID No. 11); Vgl(fx) (Seq. ID No. 12); Vgr-1(fx) (Seq. ID No. 13); and GDF-1(fx) (Seq. ID No. 14).

10

30. A method for inducing hepatic tissue formation at a damaged tissue locus in a mammalian liver comprising providing to said locus a therapeutic amount of a morphogen comprising at least  
15 residues 43-139 of hOP-1 (Seq. ID No. 5).

31. A method for diagnosing tissue dysfunction in a human, the method comprising the steps of :  
20 (a) repeating, at intervals, the step of detecting the concentration of endogenous anti-morphogen antibody present in a human; and  
(b) comparing said detected concentrations, wherein changes in the detected concentrations are  
25 indicative of status of said tissue.

32. A method for evaluating the status of a tissue, the method comprising the step of detecting the concentration of a morphogen present in said  
30 tissue.

33. The method of claim 32 comprising the additional steps of:  
35 (a) repeating, at intervals, the step of detecting the concentration of morphogen present in said tissue; and

(b) comparing said detected concentrations, wherein changes in said detected concentrations are indicative of the status of said tissue.

5

34. The method of claim 33 wherein said morphogen is selected from the group consisting of: hOP1 (Seq. ID No. 5); mOP1 (Seq. ID No. 6); hOP2 (Seq. ID No. 7); mOP2 (Seq. ID No. 8); CBMP2A(fx) (Seq. ID No. 9); CBMP2B(fx) (Seq. ID No. 10); DPP(fx) (Seq. ID No. 11); Vgl(fx) (Seq. ID No. 12); Vgr-1(fx) (Seq. ID No. 13); and GDF-1(fx) (Seq. ID No. 14).

35. A morphogen useful in the manufacture of a pharmaceutical for use in the induction of non-chondrogenic mammalian tissue growth.

36. A morphogen useful in the manufacture of a pharmaceutical for use as an inducer of progenitor cell proliferation.

37. A morphogen useful in the manufacture of a pharmaceutical for use in maintaining the phenotypic expression of differentiated cells in a mammal.

38. A morphogen useful in the manufacture of a pharmaceutical for use in the induction of hepatic tissue growth.

30

39. The morphogen of claims 35, 36, 37, or 38 wherein said morphogen comprises an amino acid sequence sharing at least 70% homology with a sequence selected from the group consisting of: hOP1 (Seq. ID No. 5); mOP1 (Seq. ID No. 6); hOP2 (Seq. ID

35



No. 7); mOP2 (Seq. ID No. 8); CBMP2A(fx) (Seq. ID No. 9); CBMP2B(fx) (Seq. ID No. 10); DPP(fx) (Seq. ID No. 11); Vgl(fx) (Seq. ID No. 12); Vgr-1(fx) (Seq. ID No. 13); and GDF-1(fx) (Seq. ID No. 14).

5

40. The morphogen of Claim 39 wherein said morphogen comprises an amino acid sequence sharing at least 80% homology with one of the sequences selected from said group.

10

41. A morphogen useful in the manufacture of a pharmaceutical to inhibit neoplastic cell growth.

42. A cancer therapeutic agent comprising a morphogen.

15

43. A therapeutic agent for tissue growth induction, the therapeutic agent comprising a morphogen.

20

44. A therapeutic agent for inducing phenotypic expression of differentiated cells, the therapeutic agent comprising a morphogen.

45. A therapeutic agent for inducing progenitor cell proliferation, the therapeutic agent comprising a morphogen.

25

30

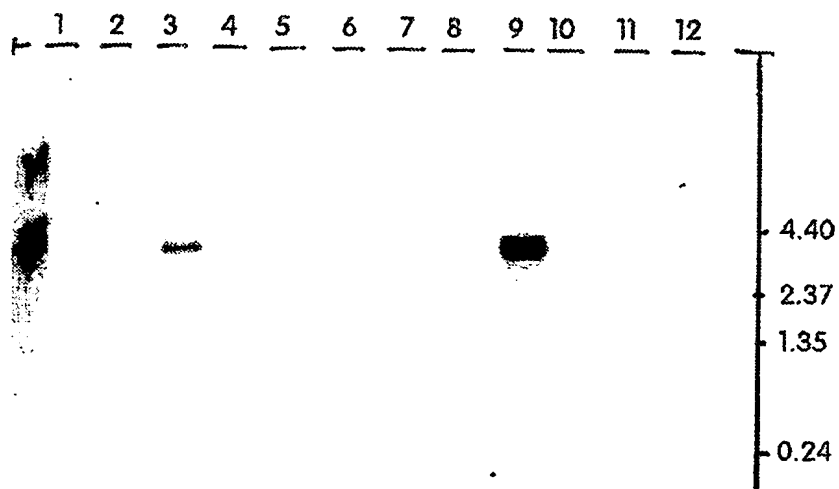


Fig. 1

2/11

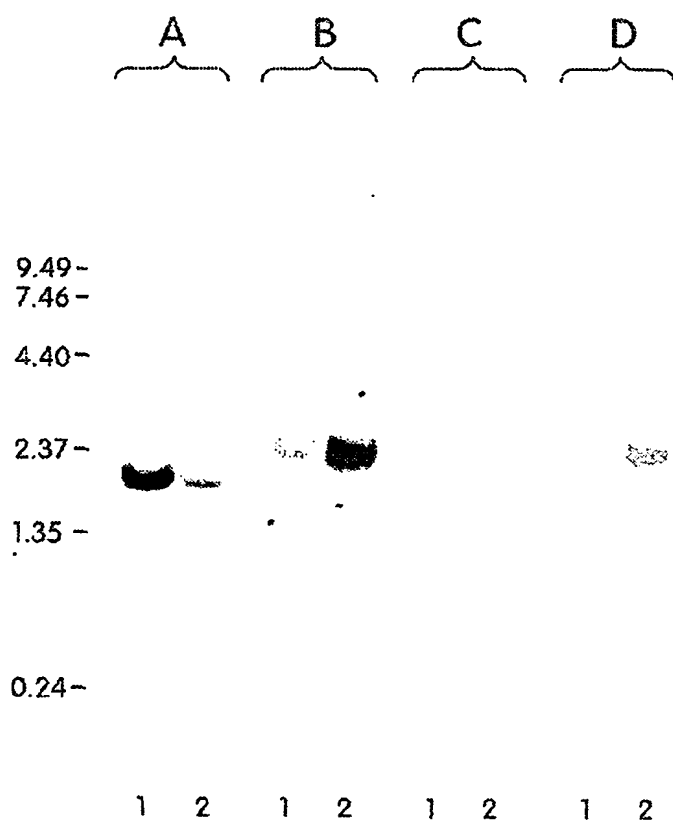


Fig. 2

3/11

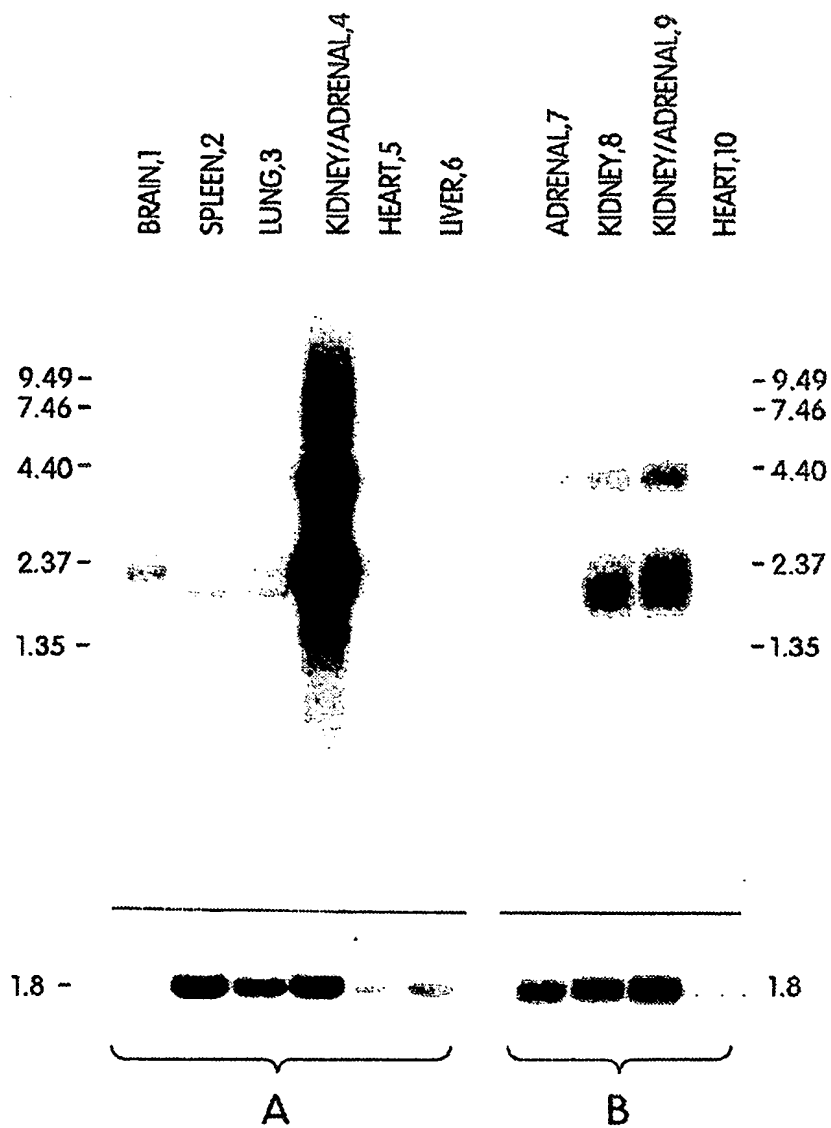


Fig. 3

4/11



Fig. 4A

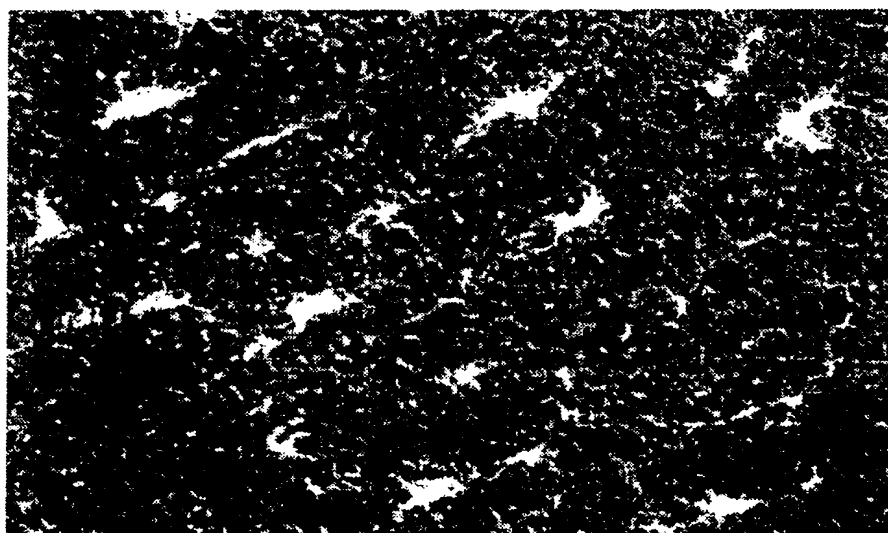


Fig. 4B

5/11



Fig. 5A

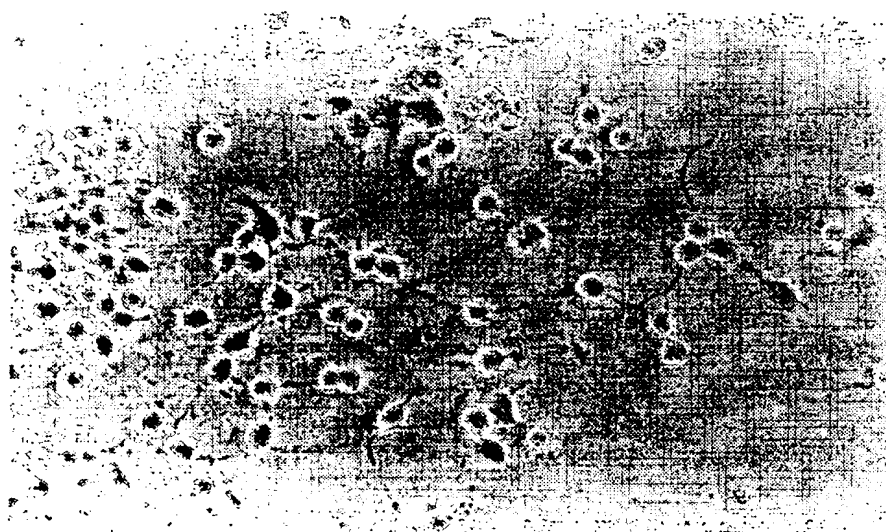


Fig. 5B

**SUBSTITUTE SHEET**

6/11

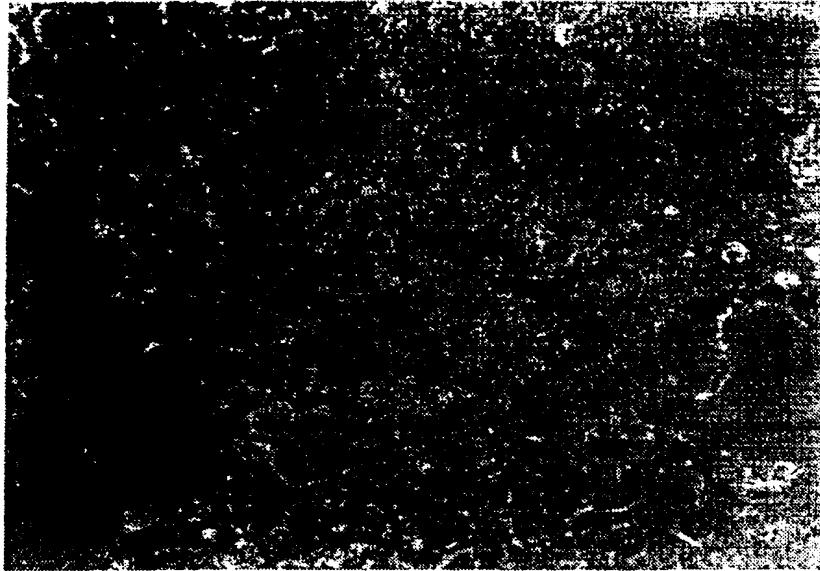


Fig. 6A



Fig. 6B

**SUBSTITUTE SHEET**

7/11



Fig. 6C



Fig. 6D

**SUBSTITUTE SHEET**



8/11



Fig. 7

SUBSTITUTE SHEET

9/11

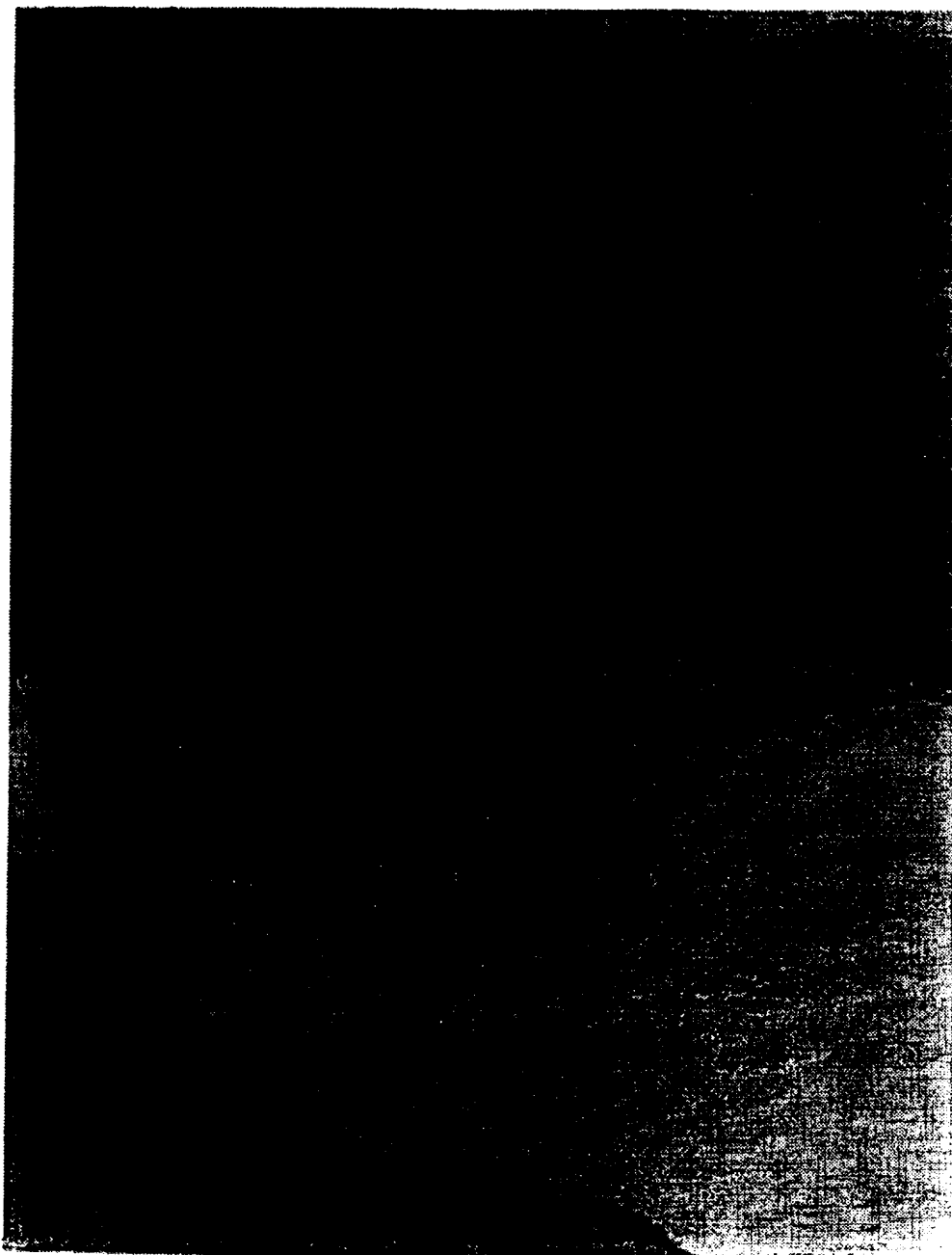


Fig. 8A

10/11



Fig. 8B

SUBSTITUTE SHEET

11/11



Fig. 8C

SUBSTITUTE SHEET

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US92/01968

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> (If several classification symbols apply, indicate all) <sup>3</sup>		
According to International Patent Classification (IPC) or to both National Classification and IPC		
IPC (5): A61K 37/12; A61P 2/02; C07K 13/00		
US CL : 350/356, 402; 424/423, 426; 435/240.243		
<b>II. FIELDS SEARCHED</b>		
Minimum Documentation Searched <sup>4</sup>		
Classification System	Classification Symbols	
U.S.	350/356, 402; 424/423, 426; 435/240.243	
Documentation Searched other than Minimum Documentation to the extent that such Documents are included in the Fields Searched <sup>5</sup>		
CHEMICAL ABSTRACTS, APS		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <sup>14</sup>		
Category <sup>6</sup>	Citation of Document <sup>16</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
X/Y	WO, A, 89/09788 (OPPERMANN ET AL.) 19 OCTOBER 1989, see entire document.	1/5-45
X/Y	WO, A, 89/09787 (KUBERASAMPATH ET AL.) 19 OCTOBER 1989, see entire document.	1/5-45
<p>* Special categories of cited documents:<sup>16</sup></p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search <sup>4</sup>		Date of Mailing of this International Search Report <sup>2</sup>
12 June 1992		23 JUN 1992
International Searching Authority <sup>1</sup>		Signature of Authorized Officer <sup>20</sup>
ISA/US		JAMES KETTER